Virtual Reality (2025) 29:49
https://doi.org/10.1007/s10055-025-01129-w

ORIGINAL ARTICLE

Embodiment in virtual reality: an experiment on how visual and aural First- and Third-person modes affect embodiment and mindfulness

Isak de Villiers Bosman[1,2] · Mila Bujić[1] · Laura Diana Cosio[1] · Oğuz 'Oz' Buruk[1] · Kristine Jørgensen[3] · Juho Hamari[1]

Received: 14 August 2024 / Accepted: 4 March 2025

Abstract

Immersion as an umbrella experience and embodiment are regarded as primary advantages of virtual reality and its capability to lead to other reflective experiences. Beyond the basis of stereoscopic head-mounted displays, however, there is paucity of research on altered ways of experiencing virtual reality with regards to the possibilities afforded by this medium. To address this gap, we explore the effects of perspective shifts, both in terms of the visual and aural modalities, on embodiment and embodied mindfulness. Using a virtual reality application that administers a body-scan meditation using different visual- and aural perspectives, we investigate the effects these perspectives have on embodiment and embodied mindfulness. Our results indicate that a first-person visual perspective positively influenced a change in perceived body schema, but not in acceptance of virtual body ownership or control/agency of a virtual body, nor do they suggest an effect of audio-based perspective on these outcomes. Investigation into the moderating effects of mindfulness- and immersive tendencies on these factors suggest that participants with low immersive tendencies experienced a greater change in their body schema in the visual first-person condition compared to the third-person condition. A qualitative content analysis on participants' experiences indicate that few participants were able to express their experience in terms of audio. Our results contribute to the body of work on altered self-representations for mindfulness and extends on this concept with the notion of listening perspective as well as more generally to the design of sound, perspective, and embodiment in virtual reality.

Keywords Virtual reality · Audio · Perspective · Third-person · Embodiment · Mindfulness

1 Introduction

The topic of virtual embodiment has received significant attention in the last few years (Chen et al. 2023) and is increasingly emphasized as a key consideration in the design of virtual reality (VR) applications (Guy et al. 2023). Embodiment refers to the general notion of owning a body and can be conceived in various terms, including the phenomenological aspects (de Vignemont 2011). This complex topic has been studied using multiple approaches, many of which are facilitated by the unique capabilities of VR technology for representing and controlling a virtual body in the form of an avatar. One such approach is to use the ability of VR to remove the user's visual frame of reference from their avatar to create a third-person perspective (3PP) as opposed to a first-person perspective (1PP). While 3PP is common in traditional digital games, the accurate visual-motor synchrony provided by VR technology creates new opportunities for designing such experiences and a 3PP has been found to provide advantages in VR, such as an improved sense of space (Gorisse et al. 2017) and a reduced sense of motion sickness (Medina et al. 2008; Monteiro et al. 2018). However, the use of 3PP is underutilized in embodiment research in VR (Mottelson et al. 2023) and VR in general (Black 2017; Hoppe et al. 2022). Furthermore, current research into the use of 3PP in VR applications focuses on the visual modality with little to no emphasis given to the audio modality. There are several reasons why audio feedback is relevant in the context of embodiment and altered

✉ Isak de Villiers Bosman
isak.bosman@up.ac.za

1 Gamification Group, Tampere University, Tampere, Finland

2 Department of Information Science, University of Pretoria, Pretoria, South Africa

3 Department of Information Science and Media Studies, University of Bergen, Bergen, Norway

Published online: 14 March 2025

perspectives. Audio feedback can be used to discern rich and detailed spatial qualities which is further enabled by the head-tracking provided by VR headsets (Jenny and Reuter 2021), which makes it potentially valuable as a means of providing/modifying perspectival information. These qualities also position audio as a potentially fruitful avenue for exploring embodiment-related outcomes, since spatial listening is affected by individual anatomy (Foglia and Wilson 2013) and could provide a means of diversifying design considerations for embodied experiences in VR.

The purpose of this study was to investigate the effects of a 1PP and 3PP on the experience of embodying a virtual avatar and on how the avatar affects the user's sense of their own body ownership. Furthermore, we introduce the concept of aural perspective as distinct from visual perspective and investigate its effect separately, hence our study uses a shift in these perspectives as the basis for investigating the malleability of embodiment aspects. We also use the concept of embodied mindfulness as an additional lens for investigating embodiment phenomena in VR. The research aims to answer the following questions:

- RQ1: How do visual and aural perspectives affect virtual embodiment in VR?
- RQ2: How do visual and aural perspectives affect embodied mindfulness in VR?
- RQ3: How does immersive tendencies moderate the effect of visual/aural perspective on virtual embodiment?
- RQ4: How does dispositional mindfulness moderate the effect of visual/aural perspectives on embodied mindfulness?

Our research contributes insight into the effects of perspective shifts on embodiment and mindfulness and introduces the concept of aural perspective as distinct from visual perspective in this context. We also explore underlying trait outcomes and their relationship with embodiment-related state outcomes as well as more general insight stemming from participants' experiences in general.

2 Background

2.1 Embodiment and virtual reality

Several approaches toward understanding the experience of using VR emphasize the role of the body, both in terms of the user and the virtual body representation (Murray and Sixsmith 1999; Southgate 2020). For example, Slater et al. (2022) consider three illusions to be key distinguishing features that are unique to VR: the illusion of being located at another location (place illusion), the illusion that depicted events are really happening (plausibility illusion), and the illusion of virtual body ownership (ownership illusion). These three illusions are generally used to describe the subjective experience of presence, with the first specifically describing the feeling of "being there" while using VR. These illusions are in turn enabled by the objective capabilities of the VR system, such as the display quality, design of environments and avatars, etc., which constitutes the immersion of the VR system (Slater et al. 2022). Phenomena such as the "Proteus effect" whereby an individual adapts their behavior to match their avatar (Yee and Bailenson 2007) also highlight the importance of virtual avatars and body ownership on users' behavior in virtual environments such as VR. Known embodiment illusions have also been replicated in VR, such as the seminal rubber hand illusion whereby an individual feels a convincing sense of ownership over a prosthetic hand, thus altering their sense of self (Botvinick and Cohen 1998; IJsselsteijn et al. 2006). An understanding of how users experience VR is thus intrinsically linked to a more general understanding of the mechanisms that underlie embodiment phenomena.

Embodiment broadly refers to the conceptualization of and issues relating to having a body (de Vignemont 2011) and can be considered in terms of various subcomponents such as the sense of body ownership, control/agency, and self-location (Kilteni et al. 2012a). A common approach for discussing embodiment is using the theory of embodied cognition, which challenges the notion of mind-body duality by claiming that cognition and psychology are, themselves, grounded in bodily features and that mental activity is inherently tied to the properties of an individual's body (Foglia and Wilson 2013). Accordingly, one's sense of body schema, i.e., the physical properties of one's body and how they facilitate action, is not only shaped by mental processes, but also informs and constrains such mental processes in turn (de Vignemont 2011; Won et al. 2015). This internal sense of body schema also affects perception, as one's body affects how one perceives action possibilities, i.e., affordances in one's environment. For example, detection of the spatial properties of sound rely on decoding information gleaned from one's hearing apparatus, such as the distance between one's ears (Foglia and Wilson 2013). However, the phenomenology of embodiment has also been found to be malleable, as demonstrated through the ability of individuals to incorporate non-body objects such as prosthetic body parts, tools, or geometrically simple objects such as virtual balloons into their body schema (Ma and Hommel 2015; Rolla et al. 2022). In this view, objects can be temporarily incorporated into an individual's sense of self and, in doing so, become "transparent" means of interfacing with one's environment, both in terms of perceiving and acting upon affordances. A useful distinction can thus be made here between motoric embodiment, where an embodied (non-body) object is processed as part of one's bodily motor systems, and perceptual embodiment, where it is processed as part of one's perceptual systems (de Vignemont 2011).

As mentioned, there is notable similarity between conceptualizations of VR experience and theories of embodiment. One example of this is the similarity between the notion of self-location in embodiment research (Guy et al. 2023) which is similar to the place illusion commonly-studied in VR research (Slater et al. 2022), although for the purposes of this study we consider the illusion of finding oneself in a different physical location to be more appropriately-situated as part of the place illusion and not inherently describing a body-ownership illusion (Roth et al. 2017). Another example is the transparent incorporation of objects into one's body schema aligns with the idea of VR as inducing the "illusion of non-mediation" (Lombard and Ditton 1997), i.e., the illusion that a virtual environment (VE) is perceived as if directly, without the mediation of technology. This illusion also extends to embodiment phenomena, as there has been ample research indicating that VR can induce a temporary modification of a user's body schema, such as including an extra arm to perform motor tasks (Won et al. 2015), remapping the size or movement of a user's limb (Kilteni et al. 2012b; Won et al. 2015), or inhabiting a completely different body (Petkova et al. 2011). Furthermore, while some research has indicated that users more readily accept a realistic avatar as their own self (Latoschik et al. 2017) and this it induces higher levels of presence (Zhang et al. 2020) other studies have also found no effect of avatar realism on illusions of virtual body ownership (Lugrin et al. 2015a, b, c; Lugrin et al. 2015b). This supports the notion that users are able to adapt to various virtual representations regardless of realism. Striving for perfect realism in avatar design also risks inducing the phenomenon known as the "uncanny valley effect" while also requiring more time and resources to develop (Latoschik et al. 2017; Lugrin et al. 2015a, b, c).

With consideration to the intrinsic connection between embodiment phenomena and cognition, Khoury et al. (2017) formulate this connection in terms of "embodied mindfulness". This notion is similar to other conceptualizations of mindfulness in that it emphasizes a detached meta-awareness of and non-reactivity to thoughts and emotions, sometimes referred to as "decentering", toward an intentional self-regulation process (Bernstein et al. 2015; Zhang et al. 2021), which is practiced through various forms of meditation (Millière et al. 2018). This ability to "step back" and observe one's own emotions and experiences as if from a distanced or "third-person" perspective as well as a general awareness of one's own body both aid in fostering emotion regulation which leads to many of the benefits conferred by mindfulness (Zhang et al. 2021). Where embodied mindfulness differs from other conceptualizations is in adopting an embodied cognition approach, whereby attention to bodily states and the mind-body connection is emphasized as a skill that promotes emotional self-regulation (Khoury et al. 2017, 2021). For the purpose of this study, this notion of mindful awareness that is grounded in bodily states was considered a useful lens for understanding embodied experiences, as this type of interoceptive-focused approach is also under-utilized in the literature on VR in mindfulness (Arpaia et al. 2022).

Perceptual embodiment can also extend to senses other than sight. For example, sound perception is affected by properties of individual human anatomy, such as the distance between one's ears (Foglia and Wilson 2013). VR technology affords embodied listening through the headphones and head-tracking provided by the technology to create the illusion of perceiving sound in 3D space, i.e. spatial audio. This is accomplished in different ways, such as by modeling aural transformations that would normally result from the way sound interacts with one's physical anatomy. These include two processes: firstly, the time and volume of the signal respectively sent to the left and right headphone are offset based on the distance between one's ears, which are referred to as interaural time and level differences (ITDs and ILDs). Secondly, the spectral transformations that would normally result from a sound wave's reflection off a person's anatomy are modeled virtually, which is referred to as head-related transfer functions (HRTFs) (Jenny and Reuter 2021). The decoding of these properties by the brain allows the listener to localize the position of sound sources in 3D space, i.e., to pinpoint their position in 3D space, whether physical or illusory, although this is limited by the fact that these characteristics are specific to an individual's anatomy, thus limiting the accuracy with this can be done using generic models for, e.g., HRTFs (Paterson and Kadel 2023). Head tracking in terms of rotation and movement also affords exploration of sonic environments, as this spatial property allows both the location of the 3D position of sound sources and the use of this information to discern one's own location in relation to the environment. Since the spatial properties in VR are algorithmically generated, they can be manipulated, which means that a virtual listening experience that is facilitated by the perceptual and motoric affordances of VR can thus be a valuable tool for understanding different aspects of VR experience more generally, including embodiment and place illusion. However, within VR research, there is a limited focus on the effects of audio feedback on embodiment-related outcomes, such as body ownership (Bosman et al. 2023). One approach for exploring embodiment in VR, especially perceptual embodiment, is using the established notions of first- and third-person perspectives (1PP and 3PP) and focusing separately on the visual and aural senses.

2.2 Perspective in VR

Headset-based VR provides intuitive support for interaction with a virtual environment through natural movements such as moving one's head and hands around in space. As such, VR experiences are overwhelmingly created with the assumption of experiencing and interacting with the virtual environment as if embodying a virtual avatar from a 1PP (Black 2017; Hoppe et al. 2022). This is in contrast with a 3PP, common in many digital games, where a user controls an avatar that is located some distance away from the virtual camera and where the user thus interacts with the virtual environment through a separate entity (Jørgensen 2013, pp. 129–130). Referring back to a phenomenological view of embodiment, the control of a 3PP avatar can be conceived in the same way as the incorporation of an external object/tool into one's body schema as VR technology provides the means of creating an accurate synchrony between physical movement and avatar behavior (Debarba et al. 2015; González-Franco et al. 2010).

On the one hand, the assumption of a 1PP in VR makes sense insofar as this is our default mode of existence in the physical world and that our default perception of a body in 1PP is that of our own (Guy et al. 2023; Won and Zhou 2024). As such, a 1PP with visuo-motor synchrony between user and avatar movement[1] most intuitively exploits the capabilities of VR in transparently enabling the integration of a virtual avatar into one's body schema, which has been found to lead to an increased sense of body ownership (Cui and Mousas 2023; Gorisse et al. 2017) and a closer match between avatar and user movement compared to 3PP (Won and Zhou 2024). A 1PP has also generally been found to improve the sense of agency over a virtual avatar (Mottelson et al. 2023), although there have been findings that contradict this, e.g. (Gorisse et al. 2017). Furthermore, investigation into several other outcomes has revealed cases where a 1PP does not offer clear advantages. For example, previous studies have found mixed/contradicting effects of perspective on task performance (Debarba et al. 2017; Gorisse et al. 2017), and spatial presence (Gorisse et al. 2017; Hoppe et al. 2022). A 3PP in VR has also been shown to provide an improved sense of overall space (Gorisse et al. 2017), a sense of safety (Debarba et al. 2017), and to reduce the sense of motion sickness (Medina et al. 2008; Monteiro et al. 2018). There are also some commercial VR games that use a 3PP to create a different experience than most VR games, such as Lucky's Tale (Playful Studios 2024) and Moss (Polyarc Games 2024). However, due to the implicit assumption of a 1PP in VR, most VR games and applications are created with a 1PP (Deng et al. 2024) and there is still an overall lack of research on using different perspectives in VR (Cui and Mousas 2023; Hoppe et al. 2022).

Apart from the benefits of these and other outcomes, further exploration of perspective as a concept in VR provides a general means of diversifying design approaches for VR applications and a perceptual lens through which to reconsider the created experience. For example, there has been much research on avatar representations in VR, e.g., (Beyea et al. 2022; Bujić et al. 2023), and in digital environments more generally, e.g., (Yee and Bailenson 2007). However, if VR assumes a 1PP, alternative approaches often need to be taken for users to be able to perceive their own avatars, such as using mirrors (Bujić et al. 2023; González-Franco et al. 2010). Furthermore, since these types of approaches are relatively uncommon in real-world VR applications such as games, the salience and effect of avatar design might be less apparent in VR as a result (Deng et al. 2024). *A third-person perspective thus provides an alternative approach for users to engage with and give more focus to the design of their (potentially modified) virtual selves.*

Furthermore, we distinguish between visual and aural perspectives in VR. Visual perspective, in this case, refers to the general formulation of an in-game camera position, i.e., the point-of-view (PoV), relative to the avatar. Aural perspective in turn refers to the point from which audio feedback is heard, i.e., the point-of-audition (PoA). We chose this approach for several reasons: firstly, the notion of aural perspective, to the best of our knowledge, has not been studied in VR before. Since the implementation of a 3PP application/game requires deciding where to place the PoA relative to the PoV, our aim was to inform guidelines on this design decision from empirical data. Secondly, while a 3PP implies that the PoV must be located some distance away from the avatar (otherwise it would be 1PP), a PoA can be modified independently, which makes it a potentially useful parameter towards understanding perceptual embodiment in 3PP VR (Fig. 1).

Previous research suggests that a 1PP leads to an increased sense of body ownership as well as agency over a 3PP (Mottelson et al. 2023). We thus hypothesized that this would apply to a visual 1PP in our study as well. We further hypothesized that this would apply separately to an aural 1PP over an aural 3PP, both of which assume a visual 3PP in our study. As such, RQ1 was answered by the following hypotheses for the confirmatory analysis:

[1] While we assume the existence of a visual avatar for our work, previous results suggest that "embodying" an invisible avatar does not affect either body ownership or agency (Mottelson et al. 2023).

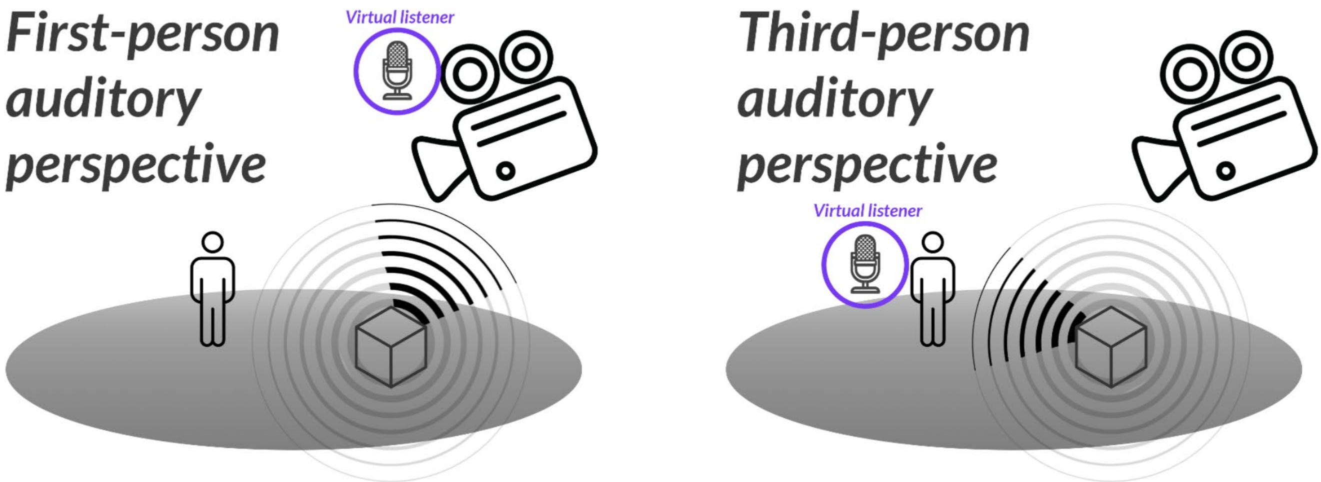

Fig. 1 A depiction of an aural 1PP (left) and an aural 3PP (right). The in-game camera is always tied to the position and rotation of the user's head. In the aural 1PP, the PoA (depicted by the microphone) is located and rotated with the camera, while in the 3PP it is done with the avatar

- H1a: The use of a visual 1PP in a VR mindfulness application will lead to an increased sense of embodiment compared to a visual 3PP.
- H1b. The use of an aural 1PP combined with a visual 3PP in a VR mindfulness application will lead to an increased sense of embodiment compared to an aural 3PP combined with a visual 3PP perspective.

In addition to this confirmatory approach, we also explored the effect of visual and aural perspectives on embodied mindfulness more broadly (RQ2) as well as moderation effects of related traits on the relationship between the conditions and outcomes in terms of virtual embodiment (RQ3) and embodied mindfulness (RQ4). To minimize spurious artefacts that might lead to false positives, we only investigated one trait outcome as a potential moderating variable for one state outcome. As such, mindful attention awareness scale (MAAS) outcomes were compared with embodied mindfulness questionnaire (EMQ) outcomes, since both pertain directly to mindfulness, and immersive tendencies questionnaire (ITQ) outcomes were compared with the virtual embodiment questionnaire (VEQ) to investigate the tendency to experience presence on different aspects of embodiment (Fig. 2).

3 Methods

Our study employed a mixed-methods approach, combining a repeated-measures within-subject experiment with a short interview (N=52). The goal of the experiment was to compare participants' experience in terms of embodiment and embodied mindfulness across three conditions: audio-visual 1PP, visual 3PP and aural 1PP, and visual 3PP and aural 3PP.

In our experiment, the audio-visual 1PP condition acted as a control condition, since this approach is the norm in the design of VR applications. A within-subjects approach was employed since the effect between the two aural conditions was expected to be relatively small and we anticipated the carry-over effects between conditions to be small enough to be managed through counterbalancing. 50 participants in a within-subjects design would have approximately 80% power in a within-subjects design for detecting a main effect of d = 0.5 when comparing the three conditions. Short interviews investigated participants' experiences of the differences between conditions, with a focus on perspective and audio, since we expected the concept of an aural perspective in particular to be relatively unknown to participants. Data collection thus involved a pre-test survey, post-test surveys after each condition, and a post-test interview after all conditions. Numerical analysis was done using Jamovi (The Jamovi Project 2022) for all tests except non-parametric post-hoc tests, which were done with IBM SPSS. Following the principles of open and reproducible science, the full study was pre-registered at: https://osf.io/v6az5. The dataset is available at: https://doi.org/10.60686/t-fsd3893.

3.1 Participants

The study included 52 participants (33 male, 18 female, 1 undisclosed) aged 19–45 (M = 27.6, SD = 5.14) with no self-assessed problems using VR. A non-probability convenience sample was used and recruitment was done using mailing lists for Tampere University and for participation in scientific research. Sample size was determined through a combination of power estimation and study time constraints. The sample size was planned at 50 participants, although more were recruited to account for dropouts, leading to the

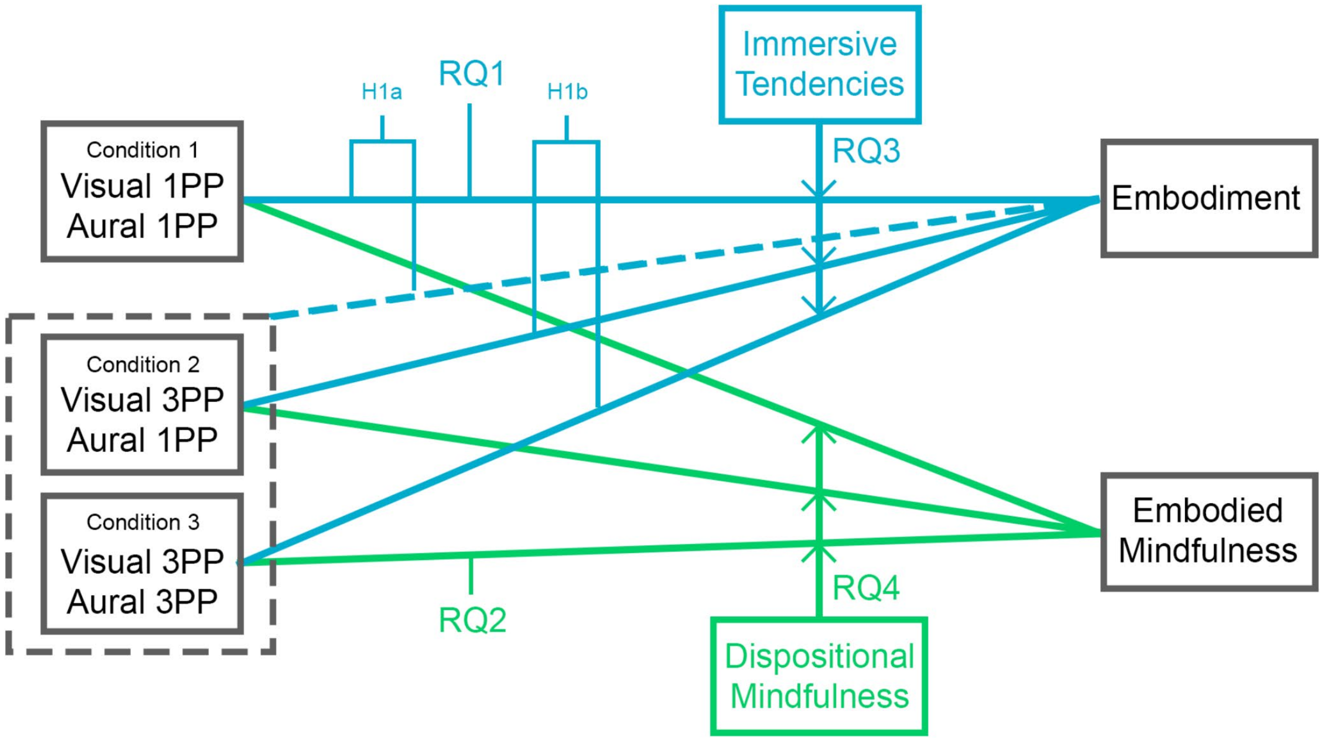

Fig. 2 The research model of the study; RQ1 is addressed through H1a and H1b and both RQ1 and RQ3 pertain to the embodiment outcomes; the independent and dependent variables for both are marked in blue. RQ2 and RQ4 both pertain to the embodied mindfulness outcomes; the independent and dependent variables for both are marked in green. RQ1 and RQ2 pertain to the effect of perspective changes on the outcomes and RQ3 and RQ4 pertain to the effect of moderating variables on the outcomes

final sample of 52. According to the national research guidelines for Tampere University, a consent form and privacy notice were communicated to participants beforehand along with clear indications for the possibility of withdrawal from the study at any given time. Participants were compensated with a movie voucher.

3.2 Materials

3.2.1 Tasks

To test the effects of different visual and aural perspectives, we developed three different versions of a VR application. We chose to use a mindfulness exercise in VR as the means to expose participants to different perspective conditions. This was done for two reasons: firstly, there was an emphasis on exposing participants to aural feedback, which required a context where the audio was central to the overall experience. A narrated mindfulness exercise was thus selected, as participants would need to listen to instructions for the duration of the experiment. Secondly, based on the formulation of embodied mindfulness given above, we expected that a form of mindfulness practice that instructs participants to pay attention to their own bodily sensations would attune participants to formulate their experience in terms of embodiment aspects. A body scan meditation was thus selected to draw the listener's attention to their own bodily sensations, which was expected to have the largest impact on the measured body-centered outcomes.

Each version of the VR application placed users in a virtual forest environment, as this is the most commonly used environment in studies investigating simulated environments' effects on mindfulness (Döllinger et al. 2021; Wang et al. 2024). This included forest ambient sounds as well as ambient music, both of which were recorded/created by the first author. The music was created to be slow and repetitive with few audio layers and synchronization with a fixed rhythm so as not to draw too much attention to itself. The inclusion of natural sounds was motivated by a desire to create a virtual space that is perceived by participants as comfortable (Eloy et al. 2023). In order for participants to orient themselves with their virtual body, they were given in-VR instructions to touch two floating blocks to their left and right. Upon doing so, the instructions vanished and two floating shapes appeared, which orbited in opposite directions around the virtual body's head from a random starting position. Apart from moving their heads around, participants were not able to affect the position/movement of these in any way. Both of these emitted 3D positional audio: the first emitted an additional layer of the ambient music being played and the second emitted instructions for the meditation exercise. The extra audio layer in the form of ambient music was added as an extra spatial audio source to make the differences in aural perspective more salient and low-volume white noise was added to increase the range of spectral content for the HRTFs to work effectively (Jenny and Reuter 2021) (Fig. 3).

For the meditation exercise, two different 10-minute audio files of narrated instructions for a body-scan meditation were used, which were developed for use in previous studies (Cropley et al. 2007; Ussher et al. 2009); both of these were used with permission from the authors. Two different audio clips were used because it was expected that participants could experience listening fatigue from listening to one 10-minute audio clip three times in a row, hence the two clips were alternated between and counterbalanced for different participants. The application was developed for HTC Vive using the built-in headphones of the headset to effectively reproduce the spatial fidelity of the audio; the Meta Quest 2, perhaps the most logical other option, was not selected because the built-in speakers have been anecdotally found not to effectively create spatial audio (Kim et al. 2022). The application itself was created using Unity3D and the Microsoft Spatializer for 3D audio (Microsoft 2023).

Each of the three versions of the application presented a different audio/visual perspective[2]. The first version presented all visual and aural feedback using 1PP (1v1a), as adopted by most current VR applications. The second version presented a visual 3PP through an avatar that mimicked the head- and hand-movements of the user which was in front of and slightly to the left of the user's camera-perspective; audio feedback was heard from the 1PP listener perspective (3v1a). The third version also presented a visual 3PP but audio was heard from the 3PP listener perspective of the avatar's head (3v3a). The avatar was slightly offset to the left to introduce a more noticeable difference between the two 3PP conditions, e.g., if the floating sphere was directly to the right of the avatar, it would be more or less in front of the camera. Figure 4 below shows the difference between the visual 1PP and 3PP and Fig. 1 illustrates the difference between the aural 1PP and 3PP. The avatar itself was designed using simplified geometrical shapes to avoid

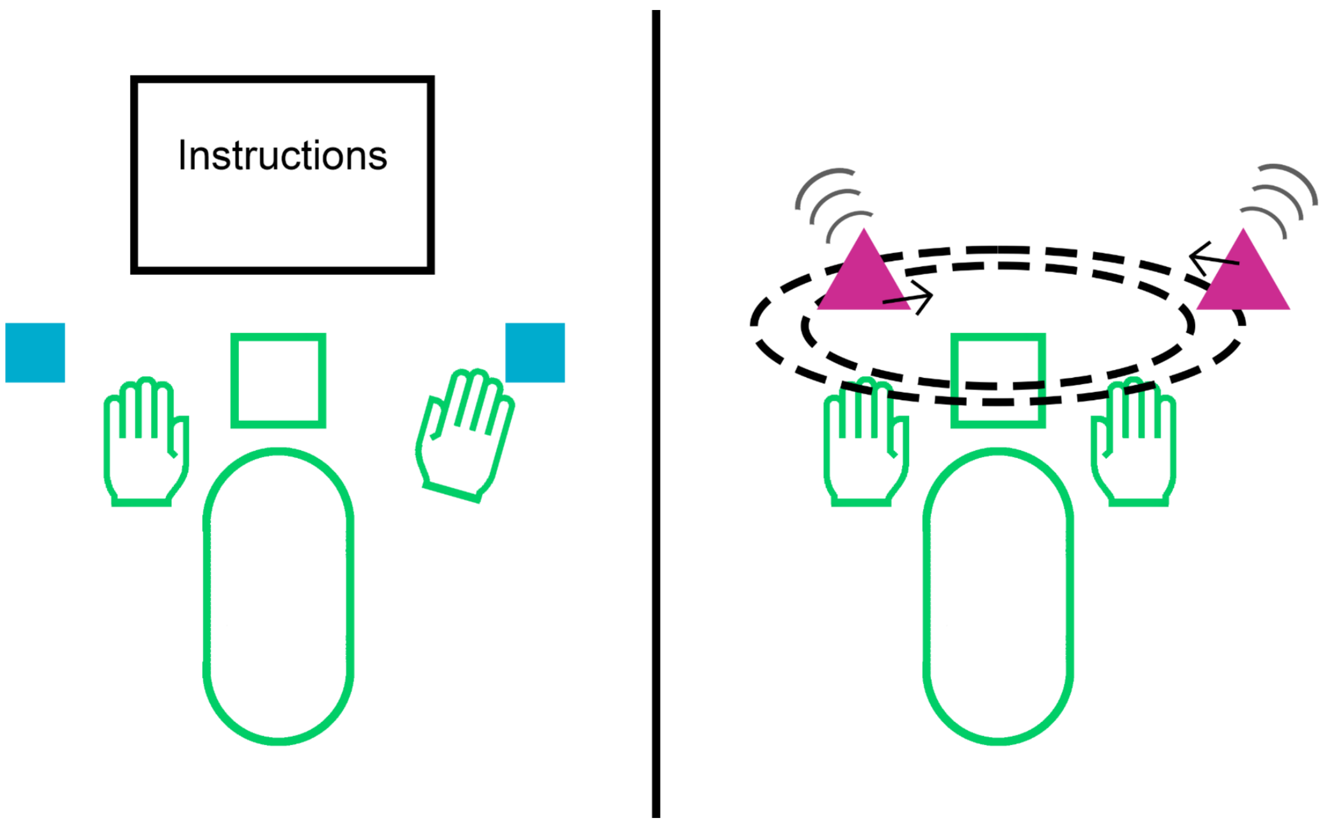

Fig. 3 The orientation process of the application used. Participants were instructed to touch two floating blocks (left) whereas two floating shapes appeared that rotated around the avatar's head and emitted 3D audio (right)

[2] While it would have been possible to combine a visual 1PP with an aural 3PP, this would involve assigning the PoA to an arbitrary point in space, which we expected participants would have difficulty in identifying without a visual point of reference.

Fig. 4 Screenshots of the VR application from the visual 1PP (left) and 3PP (right)

introducing unnecessary visual distractions from the mindfulness exercise (Döllinger et al. 2021), inadvertent coding according to specific gendered/racial characteristics, and to avoid a possible uncanny valley effect (Mori 1970). A demonstration of the application can be seen in Online Resource 1.

3.2.2 Measures

Four psychometric instruments were used, where two capturing traits were administered before the test, followed by two outcome measures after each condition. For each, the order of the items were randomized to avoid order biasing and McDonald's ω was used to measure internal consistency.

The two (pre-test) trait-based instruments used were the immersive tendencies questionnaire (ITQ; RQ3) (Witmer and Singer 1998) and the mindful attention awareness scale (MAAS; RQ4) (Brown and Ryan 2003). We also measured two (post-test) state-based outcomes with the following instruments: the virtual embodiment questionnaire (VEQ; RQ1) (Roth et al. 2017; Roth and Latoschik 2020) and the embodied mindfulness questionnaire (EMQ; RQ2) (Khoury et al. 2021, 2022).

Immersive tendencies questionnaire The ITQ describes an individual's tendency to experience presence in reaction to mediated environments and consists of 18 7-point items that pertain to an individual's reaction to various media, e.g., "*Do you easily become deeply involved in movies or TV dramas?*" from 1) Never to 7) Often. This was used to investigate possible moderating effects of these tendencies between the conditions and embodiment outcomes (RQ3). A single ITQ mean score was computed to be used in the analyses ($\omega = 0.801$).

Mindfulness attention awareness scale The MAAS describes an individual's dispositional mindfulness and is measured with 15 6-point items. Notably, all items are phrased in opposition to what is measured and are reverse coded, e.g., "*I could be experiencing some emotion and not be conscious of it until some time later*" from 1) Almost Always to 6) Almost Never; a higher mean score thus reflects a higher dispositional mindfulness score. A single MAAS mean score was computed to be used in the analyses ($\omega = 0.894$).

Virtual embodiment questionnaire The VEQ describes an individual's sense of owning and controlling a virtual body and was selected as a validated instrument for measuring virtual embodiment via three factors with three items each (7-point Likert scale from 1) Strongly disagree to 7) Strongly agree)

1. Acceptance of virtual body ownership (VEQ-ownership; 4 items): the extent to which one feels that the virtual body is a human body with parts that belong to them (e.g., "*It felt like the virtual body was my body*"). 1v1a - $\omega = 0.893$, 3v1a - $\omega = 0.926$, 3v3a - $\omega = 0.930$.
2. Control/agency of a virtual body (VEQ-control; 4 items): the extent to which one feels that they are in control of the movements of the virtual body and that these are in sync with their own movements (e.g., "*The movements of the virtual body felt like they were my movements*"). 1v1a - $\omega = 0.920$, 3v1a - $\omega = 0.908$, 3v3a - $\omega = 0.909$.
3. Change in one's perceived body schema (VEQ-change; 4 items): the extent that one feels that the size and shape

of their own body changes in response to a virtual body (e.g., "*I felt like the form or appearance of my own body had changed*"). 1v1a - ω = 0.863, 3v1a - ω = 0.924, 3v3a - ω = 0.871.

Embodied mindfulness questionnaire The EMQ was selected to provide an alternative view of mindfulness that is grounded in bodily states. The inclusion of this scale allowed for the investigation of the effect of a distanced perspective on embodiment-related outcomes that are conceptualized in terms of embodiment aspects. According to the authors' definition, embodied mindfulness should be considered as a set of skills/abilities rather than a set of traits/dispositions. The EMQ divides the general construct of embodied mindfulness into five factors (7-point Likert scale from 1) Strongly disagree to 7) Strongly agree)

1. Detachment from automatic thinking (EMQ-detachment; 5 items): the ability to regulate one's thought patterns as opposed to reactively responding to thoughts as they come up; this subscale is negatively scored (e.g., "*I get absorbed by my thoughts*"). 1v1a - ω = 0.906, 3v1a - ω = 0.852, 3v3a - ω = 0.869.
2. Attention and awareness of feelings and bodily sensations (EMQ-attention; 5 items): the ability to pay attention to and notice changes in one's own bodily sensations (e.g., "*I am able to feel sensations throughout my body*"). 1v1a - ω = 0.900, 3v1a - ω = 0.935, 3v3a - ω = 0.851.
3. Disconnection from the body (EMQ-disconnection; 5 items): a sense of removal or separation from one's own body (e.g., "*I feel detached from my body*"). 1v1a - ω = 0.949, 3v1a - ω = 0.917, 3v3a - ω = 0.923.
4. Awareness of the mind-body connection (EMQ-mind-body; 4 items): the ability to notice links between one's bodily and emotional state such as between anger and tension (e.g., "*I notice the link between feeling anxious and unease in my body*"). 1v1a ω = 0.726, 3v1a - ω = 0.824, 3v3a - ω = 0.826.
5. Acceptance (i.e. non-avoidance) of feelings and bodily sensations (EMQ-acceptance; 5 items): the tendency to accept negative or unpleasant sensations or feelings (as opposed to avoiding or distracting oneself from them); this subscale is negatively scored (e.g., "*I distract myself from unpleasant sensations*"). 1v1a - ω = 0.852, 3v1a - ω = 0.903, 3v3a - ω = 0.882.

For the purposes of this study the items of the EMQ were altered from being phrased in terms of general statements about an individual's daily experience to referring to their experience with the VR application being used for the study. For example, "*I get absorbed by my thoughts*" was rephrased as "*I was absorbed by my thoughts*". The scale for the items was also changed to be scored from 1) Strongly Disagree to 5) Strongly Agree.

After exposure to all conditions, participants were interviewed about their overall experience, whether they could tell apart the two third-person conditions, and how their experience was affected by the differences in perspective and audio. This was done to provide additional insight into their experience as a supplementary source of information for the interpretation of the questionnaire data. The relevant interview questions are included in Appendix A.

3.3 Procedure

Prior to the experiment, interested participants were provided with information about the study and asked to fill in a survey which included demographic information as well as the ITQ and MAAS. Attention checks were included in the pre-test surveys and participants with failed attention checks were asked to redo their surveys. Upon completion of the pre-test survey, they were given the details to book their time for the in-person experiment. During the experiment they were helped with the VR equipment and exposed to the three conditions. The order of the conditions was randomized both in terms of the order of the three conditions (3! = 6 orders) and the audio clips used (1-2-1 or 2-1-2), resulting in 12 (6 × 2) possible orders. Order randomization was done prior to the experiment in Microsoft Excel.

Interviews were automatically transcribed with the Microsoft Office transcription tools, after which the first author manually corrected transcriptions and replaced mention of the order of conditions with the specific condition depending on the randomized ordering for the participant. A simple content analysis (Elo and Kyngäs 2008) was performed on the interview transcripts with the goal of providing further insight into the quantitative results by highlighting relevant aspects of participants' experiences. The goal of this process was thus to support the quantitative analysis by providing further explanation of the differences between conditions rather than interpreting the interview data as if separate from the experimental conditions. The focus of this process was investigating the differences between conditions and which aspects (if any) of participants' experiences led them to perceive conditions as being different. As such, a deductive approach was taken initially which was guided by the study hypotheses' grouping of the three conditions, i.e., with H1a referring to the difference between the 1v1a and 3v1a/3v3a conditions and H1b referring to that between the 3v1a and 3v3a conditions; these groupings were used as initial high-level categories and quotes were then selected that pertain to these categories.

The next step involved an inductive coding approach (Elo and Kyngäs 2008) that aimed to identify why participants experienced different groups of conditions differently. Within the existing higher-level categories based on H1a or H1b, as described above, codes were created to summarize the contents of each quote. This process was then iterated on to merge codes that were found to be similar and create generic-level categories of these codes. This led to the division of codes pertaining either to audio or visual feedback and the final list of codes within these generic categories. A further subdivision was made for codes pertaining to the different preferences for aspects of the 1PP and 3PP perspectives in terms of visual feedback. No code was created for simply noting the fact that there was a visual perspective shift, since it was expected that each participant would have noticed this.

Code instances were counted to give some indication of the prevalence of noted difference as further explanation for the statistical differences, or lack thereof, between conditions; however, caution was taken not to attribute too much explanatory power to these numbers, especially given the overall small number of instances for most codes (Sandelowski 2001). The coding process was done by the first author and the list of codes and categories were checked by the fourth author (Fig. 5).

4 Results

The investigation into the effects of altered perspectives on embodiment and mindfulness was done using both a confirmatory and exploratory approach. A confirmatory approach was used for testing the hypotheses relating to virtual embodiment and an exploratory approach was used to investigate the effects on embodied mindfulness and the moderating effects of trait-based outcomes on state-based outcomes. The first part of this section discusses the results of the hypothesis testing focusing on virtual embodiment, followed by a more exploratory analysis of embodied mindfulness. This is followed by an investigation of possible moderating effects of traits on both virtual embodiment and embodied mindfulness. Lastly, this section includes the results of a qualitative content analysis that was performed to elaborate on the findings of the quantitative analysis. An illustration of which research questions relate to which outcomes can be found in the research model in Fig. 2.

To test the effects of the independent variables, we used a repeated-measures ANOVA to detect if there was a significant difference on the subscales of the VEQ between the three conditions, i.e., 1v1a, 3v1a, and 3v3a. The presence/absence of outliers were assessed using box plots and the test for normality was done using a Shapiro-Wilk test at a significance level of 0.05. Since none of the subscales were found to be normally distributed, a Friedman test was used as a non-parametric alternative throughout. Based on inspecting boxplots, outliers were detected and inspection revealed that one participant entered only values of 1 for each VEQ item. Cross-referencing with their interview data also indicated that their interview was much shorter than most (3:33 compared to the average of 07:23) and thus it was suspected that they entered faulty data and their entry was removed. Inspection of other outliers did not reveal such suspicions and since non-parametric tests were used, the remaining outliers were retained.

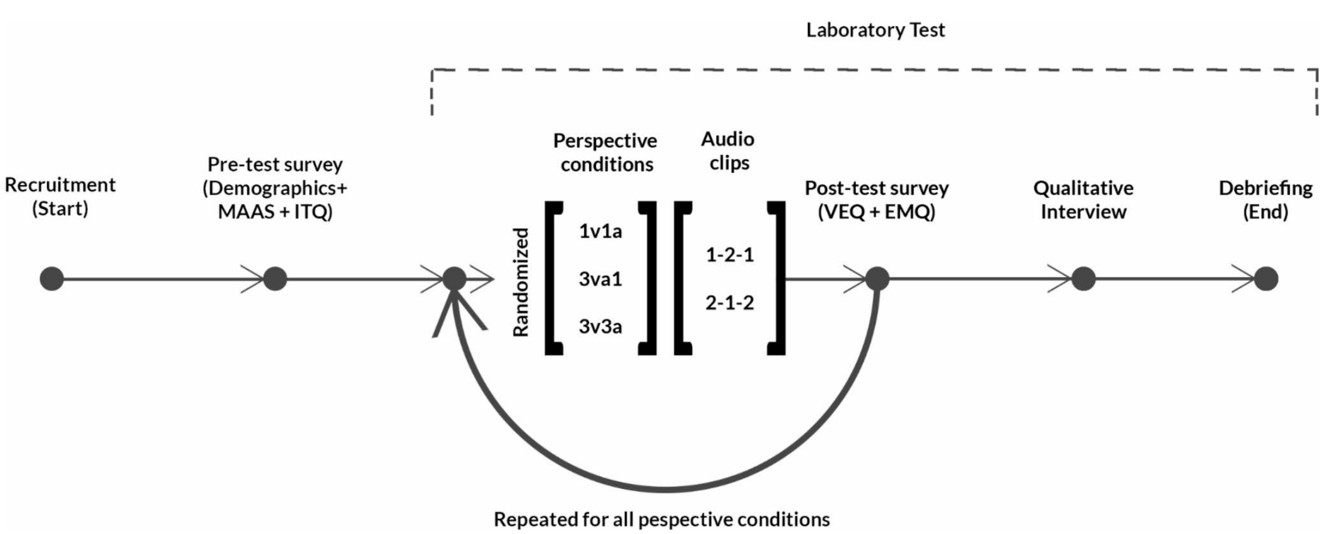

Fig. 5 The full experimental process showing the various points at which data were collected and which instruments were used

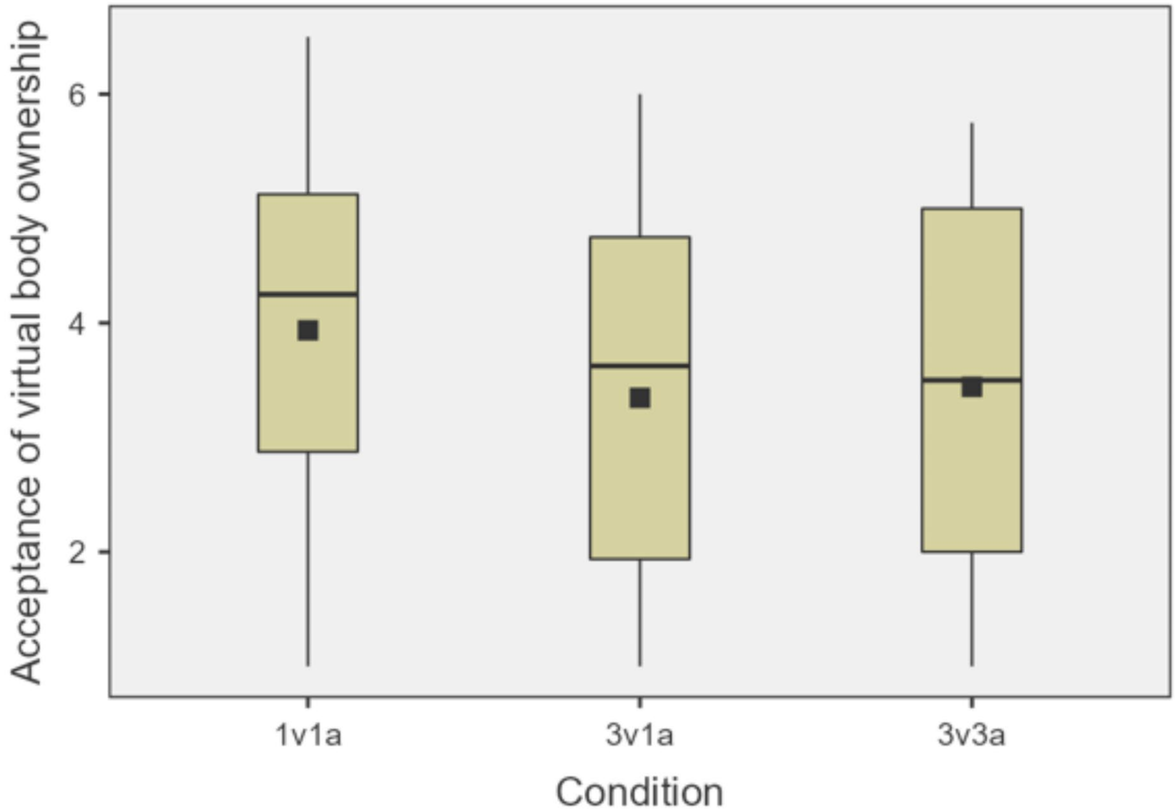

Fig. 6 Boxplots for VEQ-ownership for the three conditions

Table 1 Descriptive statistics of VEQ-ownership for the three conditions ($n=51$)

	Mean	Std. error	Median	95% confidence interval		Minimum	Maximum
				Lower	Upper		
1v1a	3.94	1.53	4.25	3.51	4.37	1.00	6.50
3v1a	3.39	1.68	3.75	2.92	3.86	1.00	6.00
3v3a	3.49	1.59	3.50	3.04	3.94	1.00	5.75

4.1 The effect of perspective on virtual embodiment

This section discusses the results of the experimental conditions on each individual subscale of the VEQ, i.e., VEQ-ownership, VEQ-control, and VEQ-change.

4.1.1 Acceptance of virtual body ownership

This factor refers to the extent to which a participant felt that the virtual body was a human body with parts that belonged to them. No outliers were detected across all three conditions, but the data were not found to be normally distributed for the 3v1a and 3v3a conditions. The Friedman test showed no significant difference between the conditions, $\chi^2(2)=5.34$, $p=0.069$. Descriptive statistics are given below (Fig. 6, Table 1).

4.1.2 Control/agency of virtual body

This factor refers to the extent to which a participant felt that they were in control of the movements of the virtual body and that these were in sync with their own movements. Outliers were detected in the data and the assumption of normality could not be restored using transformations. The Friedman test did not show any significant difference between the conditions, $\chi^2(2)=1.86$, $p=0.394$. Descriptive statistics are given below (Fig. 7, Table 2).

4.1.3 Change in perceived body schema

This factor refers to the extent to which a participant felt that the size and shape of their own body had changed while using the VR application. Although there was one outlier in the data, it was not found to have a significant effect on the results and was thus left in. The data were not found to be

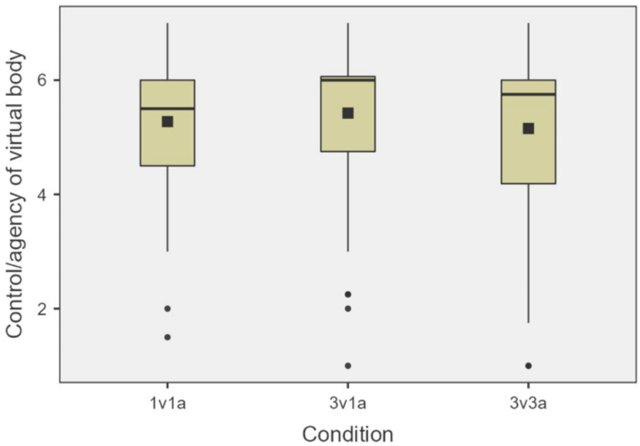

Fig. 7 Boxplots for VEQ-control for the three conditions

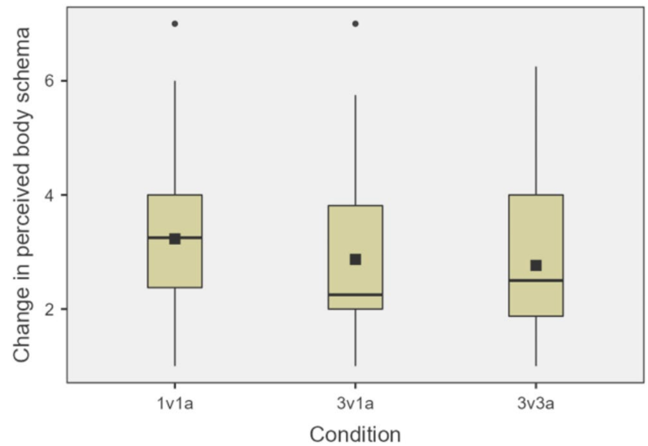

Fig. 8 Boxplots for VEQ-change for the three conditions

normally distributed and the Friedman test showed a significant effect between the conditions $\chi^2(2)=8.69$, $p=0.013$. Post-hoc testing using Kendall's W indicates a small effect size, with W=0.085, $^2(2)=8.69$, $p=0.013$ (Fig. 8, Table 3).

4.2 Exploratory analysis

4.2.1 Descriptive statistics on embodied mindfulness

Data on participants' mindfulness in relation to their bodily experience were collected using the embodied mindfulness questionnaire (EMQ). Although data were collected for each condition, we did not analyze these using inferential statistics, as we did not formulate comprehensive hypotheses pertaining to the individual subscales of the EMQ. Instead, we took an exploratory approach using descriptive statistics and boxplots to consider the data more broadly. We also used violin plots to compare the density of data points for different subscales and conditions.

Inspection of boxplots revealed several outliers for two subscales: EMQ-attention and EMQ-disconnection. However, these were not removed as they were not expected to have a notable effect on the boxplots as well as to provide an unaltered description of the EMQ data (Fig. 9, Table 4 and 5).

Inspection of descriptive statistics and box plots revealed that EMQ-attention and EMQ-disconnection have a relatively high density of points around the medians, leading to their skewness and providing explanation for outliers for both subscales. The data do not suggest notable differences between conditions on the same subscale, which accords with the general outcomes for the VEQ discussed in the previous section. The only difference where the boxplots seem to indicate some difference is for EMQ-disconnection between the 1v1a (M=2.39, Mdn=2.20) and 3v1a (M=2.28, Mdn=2.00), although this difference is small and the data are quite left-skewed. In terms of comparisons between the aural perspectives, differences tended to be quite small and several scores were more or less equal, most notably EMQ-attention and EMQ-disconnection.

4.2.2 Moderating effects of trait variables

Using the pre-test questionnaires administered to participants, i.e., the immersive tendencies questionnaire (ITQ) and mindful attention awareness scale (MAAS), we

Table 2 Descriptive statistics of VEQ-control for the three conditions ($n=51$)

	Mean	Std. error	Median	95% confidence interval		Minimum	Maximum
				Lower	Upper		
1v1a	5.27	1.23	5.50	4.93	5.62	1.50	7.00
3v1a	5.51	1.22	6.00	5.17	5.85	1.20	7.00
3v3a	5.24	1.44	5.75	4.83	5.64	1.00	7.00

Table 3 Descriptive statistics of VEQ-change for the three conditions ($n=51$)

	Mean	Std. error	Median	95% confidence interval		Minimum	Maximum
				Lower	Upper		
1v1a	3.19	1.46	3.13	2.78	3.59	1.00	7.00
3v1a	2.87	1.51	2.25	2.45	3.29	1.00	7.00
3v3a	2.76	1.41	2.50	2.37	3.16	1.00	6.25

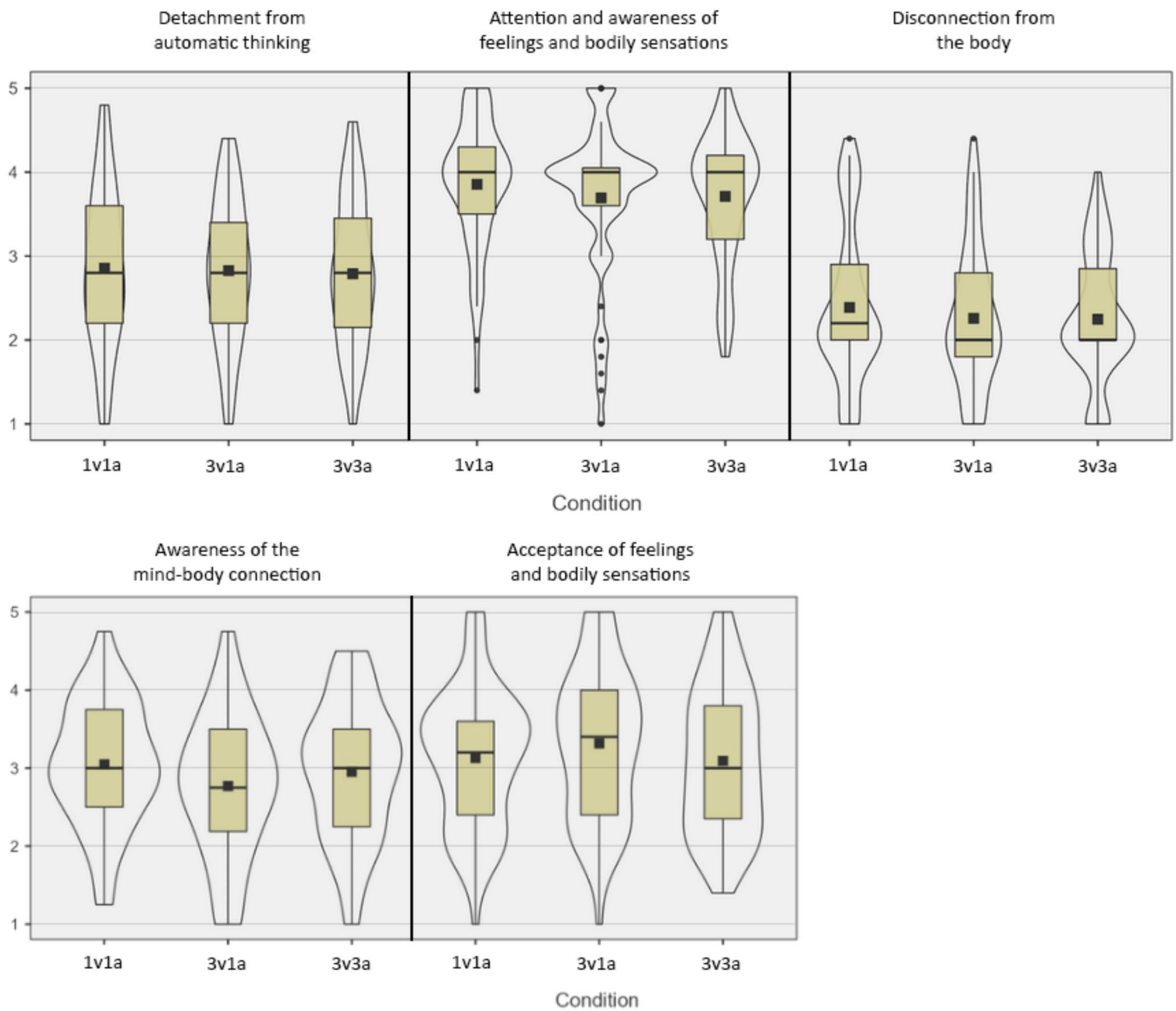

Fig. 9 Violin plots for all five EMQ subscales across all three conditions

examined the relationship between the scores for these trait-based outcomes and the post-test state-based outcomes. The purpose of this approach was to investigate possible moderation effects between state-based outcomes measured in the study and traits that are related to these outcomes. However, an exploratory approach using visual representations of data through boxplots was used over inference testing to allow for comprehensive understanding of the data with regards to its central tendency and dispersion. Each trait outcome was only matched with one state outcome to limit the number of effects being explored as a way to manage the rate of spurious artifacts leading to false positives. As such, MAAS outcomes were compared with embodied mindfulness questionnaire (EMQ) outcomes, since both pertain directly to mindfulness, and ITQ outcomes were compared with the virtual embodiment questionnaire (VEQ) to investigate the tendency to experience presence on different aspects of embodiment (as indicated in Fig. 2). For both MAAS and ITQ scores, values were transformed into binary values (low = value < median, high = value ≥ median) and the *flexplot* package in Jamovi was used to plot the values for each state outcome per trait high/low value and per condition, as provided in Figs. 10 and 11. This approach has a higher statistical power than a regression analysis for detecting possible moderation effects, although these results must be treated as preliminary indications of such effects.

Figure 10 shows the moderating effects of participants' tendencies to experience presence in reaction to mediating environments, as measured by ITQ scores, on participants' sense of virtual embodiment. *There were modest indications that a higher immersive tendency led to an increased sense of VEQ-control (in the 3v3a condition) and of larger*

Table 4 Descriptive statistics for all EMQ subscales and conditions ($n=51$)

EMQ Subscale	Condition	Mean (SD)	Median	95% CI		Min	Max
				Lower	Upper		
Detachment from automatic thinking (EMQ-detachment)	1v1a	3.15 (0.947)	3.20	2.88	3.41	1.20	5.00
	3v1a	3.16 (0.862)	3.20	2.93	3.40	1.60	5.00
	3v3a	3.24 (0.865)	3.20	2.99	3.48	1.40	5.00
Attention and awareness of feelings and bodily sensations (EMQ-attention)	1v1a	3.85 (0.785)	4.00	3.63	4.08	1.40	5.00
	3v1a	3.75 (0.853)	4.00	3.51	3.99	1.40	5.00
	3v3a	3.75 (0.715)	4.00	3.55	3.95	2.20	5.00
Disconnection from the body (EMQ-disconnection)	1v1a	2.39 (0.933)	2.20	2.13	2.65	1.00	4.40
	3v1a	2.28 (0.836)	2.00	2.05	2.52	1.00	4.40
	3v3a	2.27 (0.765)	2.00	2.06	2.49	1.00	4.00
Awareness of the mind-body connection (EMQ-mind-body)	1v1a	3.05 (0.797)	3.00	2.82	3.27	1.25	4.75
	3v1a	2.80 (0.884)	2.75	2.56	3.05	1.00	4.75
	3v3a	2.93 (0.870)	3.00	2.69	3.18	1.00	4.50
Acceptance of feelings and bodily sensations (EMQ-acceptance)	1v1a	3.13 (0.878)	3.20	2.88	3.38	1.00	5.00
	3v1a	3.28 (0.943)	3.40	3.02	3.55	1.00	5.00
	3v3a	3.08 (0.934)	3.00	2.82	3.35	1.40	5.00

VEQ-change (in the 3v1a condition) compared to a lower immersive tendency. Both of these differences, however, are relatively small and there are high levels of skewness in both subscales.

In terms of difference between conditions, *participants with a low tendency for experiencing presence in mediated environments experienced a smaller change in their perceived body schema in both visual 3PP conditions than in the visual 1PP condition.* These participants also experienced an increase in control/agency over the virtual avatar between the 1v1a and 3v1a conditions, but not between 1v1a and 3v3a; the difference for this subscale is also less notable than for the change in body schema. Boxplots do not suggest any notable differences between the conditions for participants with high ITQ scores.

The data shown in Fig. 11 suggest that individuals with higher levels of dispositional mindfulness as measured by MAAS scores also experience higher levels of detachment from automatic thinking, especially in the 3PP conditions. The same applies to levels of acceptance of feelings and bodily sensations although the relationship here is less evident and mostly reflected in the 3v3a condition. Similarly, higher levels of dispositional mindfulness seemed to lead to a lower sense of disconnection from the body. The data do not suggest any notable relationship between dispositional mindfulness and attention and awareness of bodily feelings or sensations or awareness of the mind-body connection.

In terms of how dispositional mindfulness affected differences between conditions, the only notable difference was that participants with low dispositional mindfulness seemed to experience a lower level of EMQ-acceptance in the 3v3a condition compared to the other two, although this effect is unlikely to be significant. Level of dispositional mindfulness did not seem to moderate differences between any of the other perspective changes, thus, we do not have evidence to suggest that dispositional mindfulness moderates a change in embodied mindfulness between different perspectives.

4.3 Content analysis

A qualitative content analysis was performed in support of the quantitative analyses to provide further insight into participants' experiences that relate to the hypotheses. The

Table 5 Descriptive statistics for all EMQ subscales and conditions

EMQ Subscale	Condition	Mean (SD)	Median	95% CI		Min	Max
				Lower	Upper		
Detachment from automatic thinking (EMQ-detachment)	1v1a	3.15 (0.947)	3.20	2.88	3.41	1.20	5.00
	3v1a	3.16 (0.862)	3.20	2.93	3.40	1.60	5.00
	3v3a	3.24 (0.865)	3.20	2.99	3.48	1.40	5.00
Attention and awareness of feelings and bodily sensations (EMQ-attention)	1v1a	3.85 (0.785)	4.00	3.63	4.08	1.40	5.00
	3v1a	3.75 (0.853)	4.00	3.51	3.99	1.40	5.00
	3v3a	3.75 (0.715)	4.00	3.55	3.95	2.20	5.00
Disconnection from the body (EMQ-disconnection)	1v1a	2.39 (0.933)	2.20	2.13	2.65	1.00	4.40
	3v1a	2.28 (0.836)	2.00	2.05	2.52	1.00	4.40
	3v3a	2.27 (0.765)	2.00	2.06	2.49	1.00	4.00
Awareness of the mind-body connection (EMQ-mind-body)	1v1a	3.05 (0.797)	3.00	2.82	3.27	1.25	4.75
	3v1a	2.80 (0.884)	2.75	2.56	3.05	1.00	4.75
	3v3a	2.93 (0.870)	3.00	2.69	3.18	1.00	4.50
Acceptance of feelings and bodily sensations (EMQ-acceptance)	1v1a	3.13 (0.878)	3.20	2.88	3.38	1.00	5.00
	3v1a	3.28 (0.943)	3.40	3.02	3.55	1.00	5.00
	3v3a	3.08 (0.934)	3.00	2.82	3.35	1.40	5.00

content map shown in Fig. 12 shows a conceptual relationship between the codes that resulted from the content analysis and in terms of the level of specificity, e.g., noticing a volume difference describes a very general-level observation, whereas a left-right panning difference implicitly inherently implies noticing a volume difference as well. The highest level of specificity is the reference to a shift in aural perspective, since a left-right panning difference is perhaps the most salient perceptual component of this shift in perspective. The difference, then, between noticing a shift in volume between one's left and right ears and a shift in perceived self-location is likely to be largely predicated upon participants' ability to separately conceptualize their own virtual self-location in terms of their visual and aural senses.

4.3.1 Differences between 1PP and 3PP

In terms of the aural differences between the third- and first-person conditions (H1a, quadrant a), participants noticed differences in terms of how audio panned between their left and right ears, e.g., hearing a sound source more to the side in one condition ($n=6$), general differences in volume ($n=6$), and differences in the voice/instructions given ($n=4$). More broadly, the most common differences between these conditions were described in terms of the visual difference between the 1PP and 3PP conditions (quadrant c). On the one hand, participants found the third-person avatar and its movement corresponding to their own to be a source of distraction away from the mindfulness exercise ($n=14$), especially in response to their own movements, e.g., P47: "*the body, at the beginning, it got me distracted. I was playing around, moving my head because it is very responsive and that's very satisfying when you move your head and it follows you. But I guess that that takes you out of the whole mindfulness thing*". Conversely, a smaller number of participants ($n=6$) noted that the third-person avatar provided a focal point for their attention while following along with the mindfulness instructions, e.g., P153: "*I think that the existence of the virtual body may have helped me focus on meditative state more than the [1v1a] intervention…I don't have any strong opinion but, I would say it was because I was able to keep looking at the virtual body. So, it worked as a focus point for me*". Other differences included broadly finding the third-person avatar to be weird/unpleasant

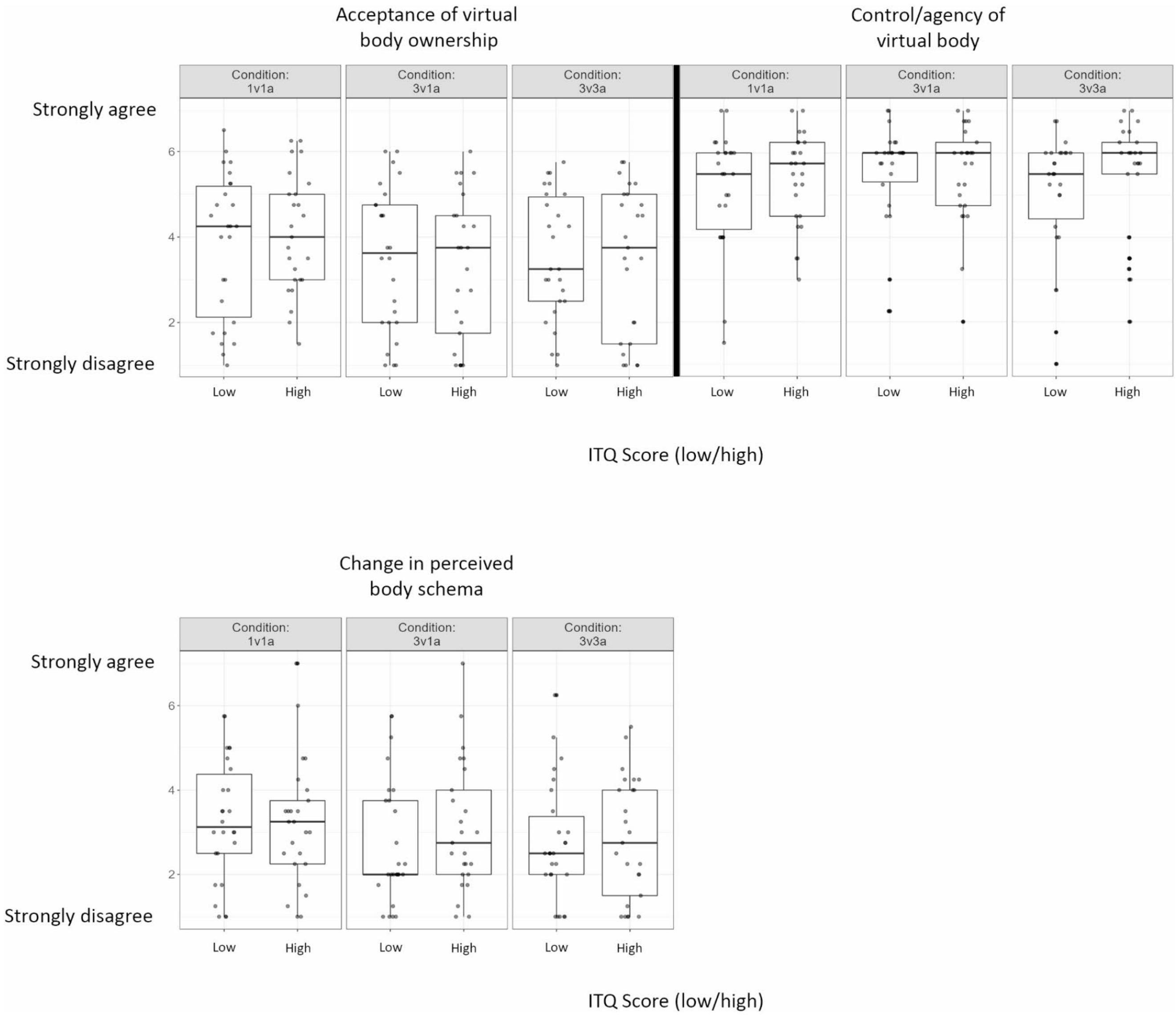

Fig. 10 Moderating effects of ITQ scores on VEQ subscales with the x-axis indicating scores for each subscale of the VEQ and the y-axis indicating low- vs. high-ITQ groups

($n=5$), and finding the visual first-person condition to provide a better sense of immersiveness/spatial presence in the environment ($n=5$) or sense of embodiment with the virtual avatar ($n=5$). The latter of these two provides partial support for the significant difference in VEQ-change between the 1PP and two 3PP conditions, although it must be noted that it could also be applied to VEQ-ownership for which no significant difference was found.

4.3.2 Differences between 3PP conditions

In terms of the difference between the 3v1a and 3v3a conditions (H1b, quadrant b), four participants described their experience in terms of a difference in aural perspective, e.g., P94: "*I think the audio was different*, *like I was in a different place compared to the source of the audio*", which most accurately describes the difference between these two conditions. While sparse, these instances provide some evidence that participants were able to separately conceive of two points from which perception is taking place across different sensory modalities, although they still described their perceived self-location in terms of their visual perspective and described their aural perspective in terms of a separate entity, e.g., "*the audio was*, *it sounded as though that person was experiencing it*, *not me*". Two more participants described the differences in sound in terms of how audio panned between their left and right ears as well as noting a general difference in volume ($n=4$) and different voices/instructions ($n=4$). The data thus suggest that being able to conceptualize a sense of self-location differently in terms of visual and aural senses was not common in participants' conceptual vocabulary in terms of their overall perception. Notably, the small prevalence of noted differences, especially pertaining to perceiving the audio from a different point in space and noticing audio panning differences explains the lack of significant differences between the 3v1a and 3v3a conditions, as most participants were not able to pinpoint concrete differences between these two conditions when prompted.

The fact that there is a higher number of code instances for the difference between the first- and third-person perspectives than between the two third-person perspectives is not unexpected, since the difference between the former is expected to be more salient. It should also be emphasized that one of the interview questions specifically probed participants about the perceived aural differences, which would have inflated the relative number of audio-centered codes. Despite this, the small number of audio-centered codes relative to visually-centered ones, including noting the different audio clips for the guided meditation, still anecdotally points toward an ocularcentrism, i.e., visual-first reflection of participants' general experiences. Furthermore, the small number of overall codes points toward a general lack of ways in which participants were able to formulate their experience with regards to specific aspects of the experience, such as the different feedback modalities or effect of the avatar which, again, supports the general lack of statistical significance in the inference tests. While it was our intention to avoid priming participants with knowledge of an audio-focused approach and specifically with the concept of an aural perspective as distinct from a visual perspective, these results suggest that the focus on such audio-centered concepts might be too niche and unfamiliar to most participants to adopt this type of design with participants who are naive to the concepts being investigated.

5 Discussion

5.1 RQ1: how do visual and aural perspectives affect virtual embodiment in VR?

5.1.1 Visual perspective and virtual embodiment

Our results support the notion that a visual 1PP led to a higher level of modification of participants' sense of their own physical bodies over 3PP, albeit with a small effect size, but that the difference in perspective did not affect the sense of ownership or control/agency over the virtual avatar. As such, our results provide only partial support for H1a. These results extend previous work on aspects of embodiment in terms of changes in perspective in VR. For example, recent work suggests that both body ownership and agency are generally affected by different perspectives (Mottelson et al. 2023; Cui and Mousas 2023), although our results do not support these findings. Our results did, however, provide support for a modified sense of body schema which has not been studied before in the context of virtual perspective-switching. However, insofar as this outcome could be considered related to the notion of responding to external stimuli as if affecting one's own body, our results contrast those of Cui and Mousas (2023) who found no effect of perspective-switching on responding to external stimuli.

There are several possible explanations for these outcomes. Regarding the lack of differences in a sense of body ownership, a possible explanation for this is that someone using a 3PP perspective, such as a player of third-person games, adopts two perspectives at the same time, i.e., their own point-of-view in the space as well as that of the avatar (Jørgensen 2013, p. 129). This is arguably more salient in VR, since a user can alter their camera view with their head movements separate from their control of an avatar; for this reason, Hoppe et al. (2022) have even argued against the notion of 1PP vs. 3PP in VR as two discrete concepts and instead consider VR perspective as a continuum, although they do distinguish between them in their results and note an increased sense of body ownership for 1PP. Although our results differ from theirs for this subscale, this argument might still explain the lack of differences in body ownership between 1PP and 3PP, as the user is always embodying the 1PP point-of-view in space, regardless of the presence of an avatar. Another explanation is that the visual-motor synchronization for all conditions was implemented to be identical, which has been found to be a key explanatory factor for this outcome (González-Franco et al. 2010). More specific to the design of our study, the lack of body ownership modification could also be explained through the simplified avatar design that did not model a full body, but only a torso, head, and hands. This could be particularly relevant to item 3 of the VEQ: "*The virtual body felt like a human body*" (Roth and Latoschik 2020), although previous research has indicated that a realistic body representation is not a requirement for a sense of body ownership and that it is possible to experience body ownership for non-corporeal objects (Ma and Hommel 2015). Finally, it is also possible that the body-focused nature of the body-scan meditation fixated participants' attention on their own bodies and sensations, thereby hindering rather than facilitating their capacity to shift the sense of body ownership onto that of an external avatar (this might also account for the high scores for EMQ-attention and low scores for EMQ-disconnection). However, this is somewhat contradicted by the fact that there

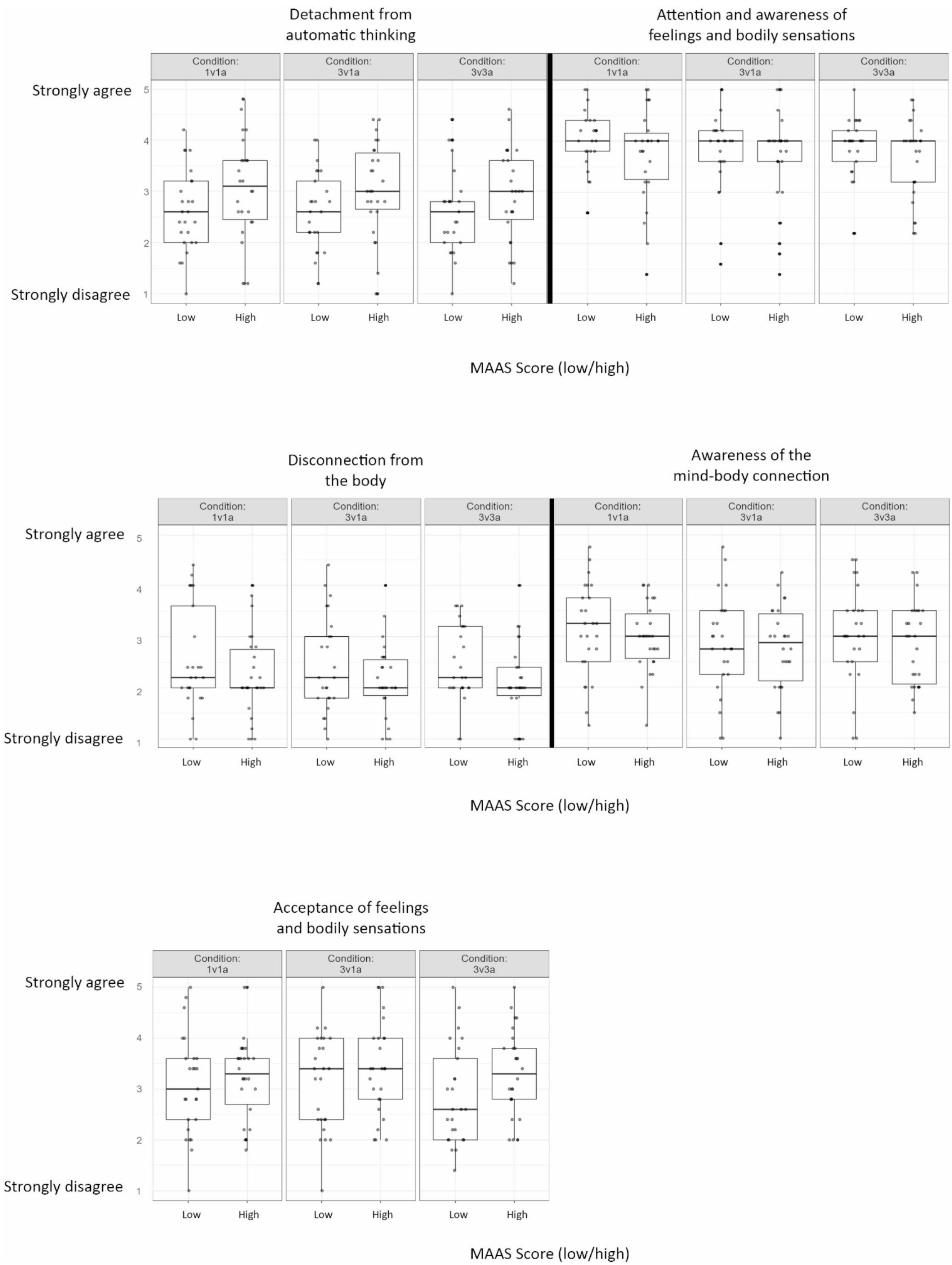
Detachment from automatic thinking
Attention and awareness of feelings and bodily sensations
Disconnection from the body
Awareness of the mind-body connection
Acceptance of feelings and bodily sensations
Condition: 1v1a
Condition: 3v1a
Condition: 3v3a
Strongly agree
Strongly disagree
1
2
3
4
5
Low
High
MAAS Score (low/high)

Fig. 11 Moderating effects of MAAS scores on EMQ subscales with the x-axis indicating scores for each subscale of the EMQ and the y-axis indicating low- vs. high-MAAS groups

was a difference in VEQ-change despite the similarity of the audio instructions across conditions.

The lack of perspective changes on the sense of control/agency of the virtual avatar is easier to account for, since the synchrony between the movements of the participant and the avatar were designed to be identical between conditions and participants reported a high sense of control/agency across all conditions (all M > 5.2, all Mdn > 5.5 out of max 7). It is also possible that the relatively simple acts of reaching out and touching blocks were not motorically demanding enough to reveal a difference in participants' sense of control over the avatar, in contrast to the more challenging game used by Hoppe et al. (2022).

We also found that a change in perspective altered participants' sense of their own body schemas, albeit with a small effect size. This is consistent with the view that users of VR are able to incorporate the technology and representations and align their sense of self to that of their virtual selves (Rolla et al. 2022). The 1PP condition was more effective than both 3PP conditions in altering participants' sense of their bodily selves, which is supported by the results of the qualitative content analysis where some participants noted a higher sense of presence, immersion, and embodiment in the 1PP condition and none noting the same for the 3PP condition.

5.1.2 Aural perspective and virtual embodiment

None of the subscales investigated were affected by a change in aural perspective and the differences between the 3v1a and 3v3a conditions were found to be minimal across all dimensions. While previous research has indicated that participants are able to discern the 3D depth of audio sources with a relatively high degree of accuracy (Turner et al. 2011), our content analysis revealed that participants were generally unable to pinpoint the difference between the two visual 3PP conditions (3v1a and 3v3a), even when specifically prompted about this difference or about changes in audio. This might be ascribed to an ocularcentrism, i.e., a focus on visual feedback which has been noted in the focus of research into various applications, such as games (Nunes and Darin 2024). Research has described a general difficulty for individuals in describing audio-focused sensations and phenomena and frequently having to resolve to using visual analogues instead (Hicks et al. 2024). As such, our results do not support H1b, nor do they support specific design implications regarding the use of aural perspective for VR in a mindfulness context, as this change did not seem to have any notable effect on participants' experiences in this regard. Further investigation into this perceptual property of spatial audio would, firstly, require emphasizing the focus on audio to participants in advance of the experimental conditions to prime them to take notice of this relatively unknown property. Secondly, it is possible that the passive nature of listening to audio instructions while being instructed to remain inactive does not allow the engagement necessary to facilitate different types of embodied listening. Considering audio perception from an embodied perspective, the VR user's body schema can be coupled to the avatar through the headset, headphones, and perceptual properties of the avatar itself (Rolla et al. 2022), as implemented in the VR application. However, this process might be contingent on audio perception as a vehicle for detecting affordances, i.e., action possibilities in an environment, in which case the ability to act upon the virtual environment would be a requirement (Millière et al. 2018). Thus, a more salient implementation of aural perspective might require a virtual environment that relies more on motoric actions and uses these to facilitate more audio-focused interaction with the sonic environment.

It is also worth pointing out that implementations for spatial audio often make use of generalized rather than individualized head-related transfer functions (HRTFs), in other words, they were created using the body-, head-, and ear anatomy of other persons. This is done for practical reasons, as creating individualized HRTFs is an expensive and time-consuming process. However, this necessarily imposes a limit on the accuracy of spatial implementations, as the decoding of HRTFs to spatial information is a highly individualized process based on embodied knowledge of one's own anatomy (Jenny and Reuter 2021). While this problem is not inherent to 1PP vs. 3PP, it is worth considering as a factor that limits the extent to which an aural perspective can be practically discerned in most VR setups using headphones. While there are other proposed solutions for VR audio that involve the use of loudspeakers and might thus reduce the need for HRTFs for spatial information (McGregor 2023), the problem of mismatched HRTFs remains a concern for the implementation of spatial audio in most commercial VR setups (Paterson and Kadel 2023). However, there has also been evidence that localization accuracy using non-individualized HRTFs improves with prolonged exposure (Stitt et al. 2019). Again, in reference to the malleable phenomenology of embodiment, continued exposure to a modified listening experience could thus allow for a modified perceptual embodiment that adopts the characteristics encoded within a different set of anatomical markers.

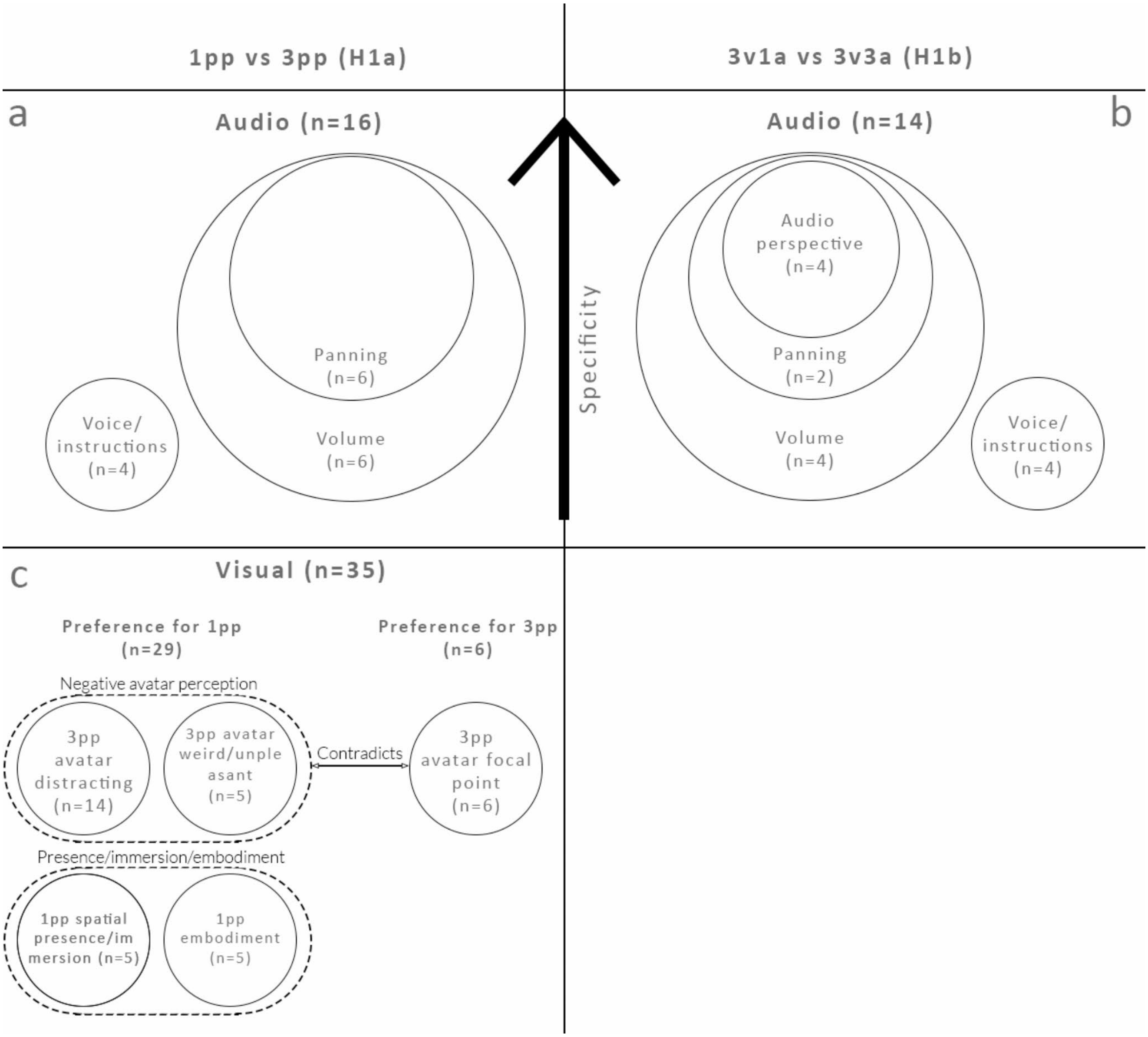

Fig. 12 A content map of the codes pertaining to the main hypotheses

5.2 RQ2: how do visual and aural perspectives affect embodied mindfulness in VR?

The exploration of perspective and embodied mindfulness continues on previous work where it was postulated that a literal representation of corporeal decentering in the form of a 3PP avatar would lead to an increased sense of decentering in a body-scan meditation (Bosman et al. 2023). This was driven by the unique ability of VR to create an altered sense of self that may be experienced in a distanced fashion and might thus alter one's sense of self (Arpaia et al. 2022). However, previous findings did not suggest any significant effects of either visual or aural 3PP on decentering and, as such, did not find support for this hypothesis. *Our exploratory analysis for the current study similarly did not suggest notable effects of perspective differences on embodied mindfulness outcomes.* As such, we do not have evidence to support the notion that any aspect of embodied mindfulness is affected by a shift in perspective in VR.

It is also noteworthy that the scores for the 2nd subscale of the EMQ are notably higher and the 3rd notably lower than the rest, considering the authors' conceptualization of the EMQ as skills that build on each other. For example, attention and awareness of bodily feelings and sensations is conceptualized as being enabled by "*Unstrapping the attention from the mind and its continuous stream of thoughts as taught in the first skill…*" (Khoury et al. 2021), i.e., detachment from automatic thinking. However, our data do not suggest a sequential structure where skills build on each other, but rather an independent building of skills.

The content analysis reveals that some participants found the visual feedback from the avatar to be a source of distraction from the mindfulness exercise, especially stemming from the responsive interactivity. Some participants did, however, find the avatar to be a useful focal point during the mindfulness exercise. This suggests that the visual capabilities of VR has the potential to aid participants in their concentration, although a balance needs to be achieved so as to not provide too much visual stimulation which might distract users from the focus of the mindfulness practice (Döllinger et al. 2021). Previous research has also suggested that the visual richness provided by VR might be detrimental rather than beneficial to mindfulness applications (Feinberg et al. 2022). Some commercially-available VR mindfulness applications, like TrippVR (https://www.tripp.com) also includes detailed, dynamic visuals and even gameplay-like functionality, such as using gaze-input to shoot colored marbles to create matches. While we consider a detailed discussion regarding the benefits and drawbacks of VR toward enabling mindfulness practice outside the scope of the current paper, a more detailed analysis of this aspect of participants' experiences can be found at [anonymized for review]. While these findings do present preliminary indications that more detailed visual feedback could be beneficial for users' focus, more research is needed to investigate the type and extent of feedback that has the potential to aid instead of hinder the mindfulness-related outcomes offered by such applications.

5.3 RQ3: how does immersive tendencies moderate the effect of visual/aural perspective on virtual embodiment?

We investigated the moderating effect of the tendency to experience presence from mediating environments, i.e., immersive tendencies, on embodiment. *Our analysis suggests that there was a slight moderation effect where participants with low immersive tendencies experienced an enhanced sense of control/agency in the 3v1a compared to the 1v1a condition, although this effect was less pronounced and not detected in the 3v3a condition.* This moderation effect could potentially be explained by the salience of the synchrony between participants' own movements and that of the avatar when they were able to see the avatar in front of them, similar to the use of mirrors to investigate the effects of synchrony in VR, e.g. (González-Franco et al. 2010). However, this explanation would not account for the difference in scores for control/agency of the virtual body between the 3v1a and 3v3a conditions for low-ITQ participants, since the visual perception of the avatar remained constant between these conditions.

The data also suggest that low-ITQ participants experienced a greater change in their own body schema in the visual 1PP condition compared to the 3PP condition. It might be possible that those who are less exposed to mediated experiences, such as digital games, are less used to 3PP avatar representations and were therefore more easily able to adapt their internal body schema to that of a 1PP avatar that more closely resembled the spatial position and movements of their physical body than the 3PP condition where this mapping was altered. Generally speaking, 3PP has also been shown to require a higher level of cognitive demand in spatial tasks (Vogeley et al. 2004), which could also account for the disparity between those with more or less exposure to different perspectives in mediated environments. As such, individuals who do not tend to experience presence in virtual environments might have more difficulty projecting their bodily selves onto an avatar and incorporating them into their own body schema (Rolla et al. 2022). However, as such an explanation would not necessarily account for the lack of similar moderation effects on virtual body ownership; further investigation into this aspect is needed with reference to the possible explanations highlighted above. The use of immersive tendencies as a main effect in future studies could thus reveal the influence this trait has on participants' ability to adapt their sense of embodiment in response to different facets of virtual body representations.

5.4 RQ4: how does dispositional mindfulness moderate the effect of visual/aural perspectives on embodied mindfulness?

Moderation effects of a dispositional mindfulness trait suggest a limited connection between this and embodied mindfulness, which is mostly limited to the EMQ subscales for detachment from automatic thinking and disconnection from the body. Individuals with higher levels of dispositional mindfulness experienced higher levels of detachment from automatic thinking and lower levels of disconnection from the body. Attention and awareness of feelings and bodily sensations as well as the two arguably "higher-level" embodied mindfulness skills, i.e., awareness of the mind-body connection and acceptance of feelings and bodily sensations were seemingly not moderated by the dispositional mindfulness trait (with the exception of the 3v3a condition for acceptance of feelings and bodily sensations). This can be explained by the formulation of embodied mindfulness and how it diverges from traditional mindfulness conceptualizations through an emphasis on the connection between the mind and body and how this can facilitate the regulation of negative emotions and sensations. Our data suggest that the first three skills that form part of embodied mindfulness align more closely with traditional mindfulness formulations while the last two diverge more notably.

The fact that individuals with higher levels of dispositional mindfulness also experienced higher levels of detachment from automatic thinking is perhaps unsurprising when comparing the similarity between the questions in this EMQ subscale and the MAAS in general, e.g., "I get absorbed by my thoughts" (EMQ) vs. "I find it difficult to stay focused on what's happening in the present" (MAAS). The same can be argued for the one modest relationship between dispositional mindfulness scores and acceptance of feelings and bodily sensations in the 3v3a condition, e.g., "I tend not to notice feelings of physical tension or discomfort until they really grab my attention" (MAAS) vs. "I distract myself from unpleasant sensations" (EMQ). Similarly, the inverse link between dispositional mindfulness and a disconnection from the body aligns with the notion of mindfulness as a moment-to-moment awareness of one's own thoughts and sensations, i.e., an increased sense of connection.

The data do not suggest any notable relationship between dispositional mindfulness and attention and awareness of feeling and bodily sensations or awareness of the mind-body connection. This can be explained in part by the formulation of mindfulness implicit in the MAAS, which does not emphasize bodily sensations to the extent that they are emphasized in the EMQ (Khoury et al. 2017). As such, the emphasis on moment-to-moment awareness of participants' tasks or environment seemingly did not transfer to their level of awareness of their bodily sensations or how their emotions manifest themselves in terms of such sensations.

Overall, the general lack of moderation effects of the MAAS score on the EMQ outcomes can be explained by the discrepancy between a conventional approach to mindfulness present in the literature and one that emphasizes the role of the body, as highlighted by (Khoury et al. 2017). This suggests that a surface-level of awareness of and attention to one's own bodily sensations, as measured by MAAS scores, does not in itself imply an ability to notice the effects of these sensations on one's own emotions or the ability to regulate them. However, as the investigation of these factors was done using an exploratory approach, they should be cautiously interpreted with consideration to the sample size and analysis approach.

6 Limitations and future work

The sample consisted mainly of students from a single university, which could skew the sample in terms of demographic factors such as age, as well as familiarity with technology and mindfulness. The virtual avatar was designed to be a simplified humanoid representation, which could have affected participants' ability to experience a sense of embodiment and familiarity with the avatar. Furthermore, the VR application was developed using the HTC Vive headset and controllers only, therefore only the head and hands could be tracked and the avatar only represented a head, body, and hands. A more sophisticated application might incorporate more detailed body-tracking to create a body representation that is more complete in terms of the mapping of different body parts, e.g., using trackers for the waist and feet. Participants were exposed to each condition with a short break in between and the order was randomized between participants; nevertheless, it is possible that there were carryover effects between the conditions from repeated exposure to the audio clips for the mindfulness exercises.

Although we were primarily interested in the embodiment aspects specifically relating to the change in perception of a virtual avatar as described in the VEQ, the inclusion of self-location as an outcome could have provided additional insight related to the spatial differences inherent in the conditions. However, an assumption that is prevalent in the construction of survey instruments used to study embodiment phenomena, including the VEQ, is that an individual is always experiencing different aspects of embodiment, such as body ownership of a single body at a time. In contrast, recent studies of the phenomenon known as heautoscopy reveals that, in very rare instances, individuals do sometimes experience an expanded sense of body ownership as well as "bilocation", i.e., an ambiguity sense of self-location that can lead an individual to believe that they are occupying multiple positions simultaneously (Szczotk and Wierzchoń 2023). Although such occurrences are associated with various disorders, such as schizophrenia, the authors do suggest VR as a tool that can be used to study these phenomena. As such, instruments to be used for future embodiment research could thus be formulated in such a way to allow for the possibility of expanded/ambiguous embodiment in the presence of external body representations. This could also be explored with the combined/alternated use of altered perspectives and mirrors in VR, thus expanding on previous mirror studies in this area.

We also relied exclusively on subjective data from participants in the form of questionnaires and interviews. Future research could include objective data such as using EEG or changes in behavior to detect changes that participants might have difficulty expressing through subjective approaches. The content analysis was performed by the first author and thus did not incorporate inter-coder reliability; this process was also guided by an a-priori process derived from the testing conditions, although a strictly inductive process might have revealed different insights. Lastly, although we purposefully avoided asking participants about aural perspective differences in a way that could prime their responses, the fact that aural perspective is generally an unfamiliar concept could have limited the extent to which participants were able to discern differences between the 3v1a and 3v3a conditions and to describe these differences. This topic can be explored more in depth through alternative implementations that specifically point participants' attention to this property. E.g., VR applications/games that not only provide audio feedback, but require participants to react to it might be better suited to studying the effects of this unfamiliar perceptual property.

7 Conclusions

To the best of our knowledge, this investigation is the first that experimentally explores the effects of aural perspectives in VR. Our investigation into virtual embodiment only revealed an effect of visual perspective on a change in perceived body schema but not on virtual body ownership or control/agency, nor in terms of aural perspective differences. We also did not find indications through exploratory analysis that these affect aspects of embodied mindfulness. An investigation of moderating effects provides preliminary evidence that individuals with low immersive tendencies might be more prone to experience a more prominent change in their own body schema in a visual 1PP than in 3PP and that high dispositional mindfulness could facilitate an increase in detachment from automatic thinking, attention and awareness of bodily feelings and sensations, and a decrease in disconnection from the body. Our results contribute to the body of work on altered self-representations for mindfulness as well as more generally to the design of sound, perspective, and embodiment in VR. We also provide suggestions for further research into aural perspective, specifically a more dedicated focus on this concept as a primer and a different context that affords more intentional sonic interaction with an environment.

Appendix

Interview questions

1. How did you experience the differences in conditions?
2. Can you identify the differences between the two third-person conditions?
3. How did the differences in perspective affect your experience of the differentversions of the application?
4. How did the differences in audio affect your experience of the different versionsof the application?

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10055-025-01129-w.

Acknowledgements This work has been supported by a UCDP grant given by the South African Department of Higher Education and Training as well as the Academy of Finland Flagship Programme [Grant No. 337653 - Forest-Human-Machine Interplay (UNITE) and the Research council of Finland [Grant No. 342144 - Postemotion].

Author contributions Conceptualization: I.d.V.B., M.B., O.B., K.J., J.H.; Formal analysis: I.d.V.B., M.B.; Funding acquisition: I.d.V.B., O.B., J.H.; Investigation: I.d.V.B., L.D.C.; Methodology: I.d.V.B., M.B., O.B.; Project administration: I.d.V.B., M.B., L.D.C., O.B.; Resources: I.d.V.B., L.D.C., O.B.; Software: I.d.V.B.; Supervision: O.B., K.J., J.H.; Validation: I.d.V.B., M.B.; Visualization: I.d.V.B.; Writing - original draft: I.d.V.B.; Writing– review & editing: I.d.V.B., M.B., L.D.C., O.B., K.J., J.H.

Funding Open access funding provided by Tampere University (including Tampere University Hospital).

Data availability Survey data that support the findings of this study have been deposited in the Finnish Social Science Data Archive at https://doi.org/10.60686/t-fsd3893.

Declarations

Competing interests The authors declare no competing interests.

References

Arpaia P, D'Errico G, De Paolis LT, Moccaldi N, Nuccetelli F (2022) A narrative review of Mindfulness-Based interventions using virtual reality. Mindfulness 13:556–571. https://doi.org/10.1007/s12671-021-01783-6

Bernstein A, Hadash Y, Lichtash Y, Tanay G, Shepherd K, Fresco DM (2015) Decentering and related constructs: A critical review and metacognitive processes model. Perspect Psychol Sci 10:599–617. https://doi.org/10.1177/1745691615594577

Beyea D, Ratan R (Robby), Lei Y (Skylar), Liu H, Hales GE, Lim C (eds) (2022) A New Meta-Analysis of the Proteus Effect: Studies in VR Find Stronger Effect Sizes. PRESENCE: Virtual and

Augmented Reality 31, 189–202. https://doi.org/10.1162/pres_a_00392
Black D (2017) Why can I see my avatar?? Embodied visual engagement in the Third-Person video game. Games Cult 12:179–199. https://doi.org/10.1177/1555412015589175
Bosman IdeV, Buruk O 'Oz,' Jørgensen, Hamari K (2023) J., The effect of audio on the experience in virtual reality: a scoping review. Behaviour & Information Technology 43, 165–199. https://doi.org/10.1080/0144929X.2022.2158371
Botvinick M, Cohen J (1998) Rubber hands 'feel' touch that eyes see. Nature 391:756–756. https://doi.org/10.1038/35784
Brown KW, Ryan RM (2003) The benefits of being present: mindfulness and its role in psychological well-being. J Personal Soc Psychol 84:822–848. https://doi.org/10.1037/0022-3514.84.4.822
Bujić M, Macey A-L, Kerous B, Belousov A, Buruk O, Hamari J, Self-representation does (not) spark joyExperiment on effects of avatar customisation and personality on emotions in VR, in: Proceedings of the 7th International GamiFIN Conference. Presented at the GamiFIN, Conference (2023) CEUR-WS, Levi, Finland, pp. 96–105
Chen J, Ning C, Chen C (2023) A systematic literature review of virtual embodied experience. Psychology & Marketing. https://doi.org/10.1002/mar.21876. n/a-n/a
Cropley M, Ussher M, Charitou E (2007) Acute effects of a guided relaxation routine (body scan) on tobacco withdrawal symptoms and cravings in abstinent smokers. Addiction 102:989–993. https://doi.org/10.1111/j.1360-0443.2007.01832.x
Cui D, Mousas C (2023) Evaluating the Sense of Embodiment through Out-of-Body Experience and Tactile Feedback, in: Proceedings of the 18th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, VRCAI '22. Association for Computing Machinery, New York, NY, USA, pp. 1–7. https://doi.org/10.1145/3574131.3574456
de Vignemont F (2011) Embodiment, ownership and disownership. Consciousness and cognition, brain and self. Bridging Gap 20:82–93. https://doi.org/10.1016/j.concog.2010.09.004
Debarba HG, Molla E, Herbelin B, Boulic R (2015) Characterizing embodied interaction in First and Third Person Perspective viewpoints, in: 2015 IEEE Symposium on 3D User Interfaces (3DUI). Presented at the 2015 IEEE Symposium on 3D User Interfaces (3DUI), pp. 67–72. https://doi.org/10.1109/3DUI.2015.7131728
Debarba HG, Bovet S, Salomon R, Blanke O, Herbelin B, Boulic R (2017) Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality. PLoS ONE 12:e0190109. https://doi.org/10.1371/journal.pone.0190109
Deng D, Buji M, Lee WC, Li MR, Hamari J, Self-avatar representation matters: Deciphering user immersion in VR games through Steam reviews, in: Proceedings of the 8th International GamiFIN Conference. Presented at the GamiFIN, Conference (2024) CEUR-WS, Ruka, Finland, pp. 132–141
Döllinger N, Wienrich C, Latoschik ME (2021) Challenges and opportunities of immersive technologies for mindfulness meditation: A systematic review. Front Virtual Real 2:644683. https://doi.org/10.3389/frvir.2021.644683
Elo S, Kyngäs H (2008) The qualitative content analysis process. J Adv Nurs 62:107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x
Eloy S, Andrade M, Dias L, Dias MS (2023) The impact of sound in People's behaviour in outdoor settings: A study using virtual reality and eye-tracking. Appl Ergon 108:103957. https://doi.org/10.1016/j.apergo.2022.103957
Feinberg RR, Lakshmi U, Golino MJ, Arriaga RI (2022) ZenVR: Design Evaluation of a Virtual Reality Learning System for Meditation, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22. Association for Computing Machinery, New York, NY, USA, pp. 1–15. https://doi.org/10.1145/3491102.3502035
Foglia L, Wilson RA (2013) Embodied cognition. Wires Cogn Sci 4:319–325. https://doi.org/10.1002/wcs.1226
González-Franco M, Pérez-Marcos D, Spanlang B, Slater M (2010) The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment, in: 2010 IEEE Virtual Reality Conference (VR). Presented at the 2010 IEEE Virtual Reality Conference (VR), pp. 111–114. https://doi.org/10.1109/VR.2010.5444805
Gorisse G, Christmann O, Amato EA, Richir S (2017) First- and Third-Person Perspectives in Immersive Virtual Environments: Presence and Performance Analysis of Embodied Users. Frontiers in Robotics and AI 4
Guy M, Normand J-M, Jeunet-Kelway C, Moreau G (2023) The sense of embodiment in virtual reality and its assessment methods. Front Virtual Real 4:1141683. https://doi.org/10.3389/frvir.2023.1141683
Hicks K, Rogers K, Gerling K, Nacke LE (2024) Juicy audio: audio designers' conceptualization of the term in video games. Proc ACM Hum -Comput Interact 8:3191–31928. https://doi.org/10.1145/3677084
Hoppe M, Baumann A, Tamunjoh PC, Machulla T-K, Woźniak PW, Schmidt A, Welsch R (2022) There Is No First- or Third-Person View in Virtual Reality: Understanding the Perspective Continuum, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22. Association for Computing Machinery, New York, NY, USA, pp. 1–13. https://doi.org/10.1145/3491102.3517447
IJsselsteijn WA, de Kort YAW, Haans A (2006) Is this my hand I see before me?? The rubber hand illusion in reality, virtual reality, and me?ixed reality. Presence: Teleoperators Virtual Environ 15:455–464. https://doi.org/10.1162/pres.15.4.455
Jenny C, Reuter C (2021) Can I trust my ears in VR? Head-related transfer functions and valuation methods with descriptive attributes in virtual reality. Int J Virtual Real 21:29–43. https://doi.org/10.20870/IJVR.2021.21.2.4831
Jørgensen K (2013) Gameworld interfaces. MIT Press
Khoury B, Knäuper B, Pagnini F, Trent N, Chiesa A, Carrière K (2017) Embodied Mindfulness Mindfulness 8:1160–1171. https://doi.org/10.1007/s12671-017-0700-7
Khoury B, Vergara RC, Sadowski I, Spinelli C (2021) Embodied mindfulness questionnaire: scale development and validation. Assessment 10731911211059856 https://doi.org/10.1177/10731911211059856
Khoury B, Vergara RC, Spinelli C, Sadowski I (2022) Embodied mindfulness questionnaire (EMQ). In: Medvedev ON, Krägeloh CU, Siegert RJ, Singh NN (eds) Handbook of assessment in mindfulness research. Springer International Publishing, Cham, pp 1–14. https://doi.org/10.1007/978-3-030-77644-2_74-1
Kilteni K, Groten R, Slater M (2012a) The sense of embodiment in virtual reality. Presence: Teleoperators Virtual Environ 21:373–387. https://doi.org/10.1162/PRES_a_00124
Kilteni K, Normand J-M, Sanchez-Vives MV, Slater M (2012b) Extending body space in immersive virtual reality: A very long arm illusion. PLoS ONE 7:e40867. https://doi.org/10.1371/journal.pone.0040867
Kim H, Choi J, Doh YY, Nam J (2022) The Melody of the Mysterious Stones: A VR Mindfulness Game Using Sound Spatialization, in: Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems, CHI EA '22. Association for Computing Machinery, New York, NY, USA, pp. 1–6. https://doi.org/10.1145/3491101.3516490
Latoschik ME, Roth D, Gall D, Achenbach J, Waltemate T, Botsch M (2017) The effect of avatar realism in immersive social virtual realities, in: Proceedings of the 23rd ACM Symposium on

Virtual Reality Software and Technology, VRST '17. Association for Computing Machinery, New York, NY, USA, pp. 1–10. https://doi.org/10.1145/3139131.3139156
Lombard M, Ditton T (1997) At the heart of it all: the concept of presence. J Computer-Mediated Communication 3:0–0. https://doi.org/10.1111/j.1083-6101.1997.tb00072.x
Lugrin J-L, Latt J, Latoschik ME (2015a) Avatar anthropomorphism and illusion of body ownership in VR, in: 2015 IEEE Virtual Reality (VR). Presented at the 2015 IEEE Virtual Reality (VR), pp. 229–230. https://doi.org/10.1109/VR.2015.7223379
Lugrin J-L, Wiedemann M, Bieberstein D, Latoschik ME (2015b) Influence of avatar realism on stressful situation in VR, in: 2015 IEEE Virtual Reality (VR). Presented at the 2015 IEEE Virtual Reality (VR), pp. 227–228. https://doi.org/10.1109/VR.2015.7223378
Lugrin J-L, Latt, Johanna L, Erich M (2015c) Anthropomorphism and illusion of virtual body ownership, in: Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments, ICAT - EGVE '15. Eurographics Association, Goslar, DEU, pp. 1–8
Ma K, Hommel B (2015) Body-ownership for actively operated non-corporeal objects. Conscious Cogn 36:75–86. https://doi.org/10.1016/j.concog.2015.06.003
McGregor I (2023) Video game audio in the metaverse and beyond. ACM Games 1:15:1–153. https://doi.org/10.1145/3593242
Medina E, Fruland R, Weghorst S (2008) Virtusphere: walking in a human size VR Hamster ball. Proc Hum Factors Ergon Soc Annual Meeting 52:2102–2106. https://doi.org/10.1177/154193120805202704
Microsoft (2023) Spatializer for Unity
Millière R, Carhart-Harris RL, Roseman L, Trautwein F-M, Berkovich-Ohana A (2018) Psychedelics, Meditation, and Self-Consciousness. Frontiers in Psychology 9
Monteiro D, Liang H-N, Xu W, Brucker M, Nanjappan V, Yue Y (2018) Evaluating enjoyment, presence, and emulator sickness in VR games based on first- and third- person viewing perspectives. Comput Animat Virtual Worlds 29:e1830. https://doi.org/10.1002/cav.1830
Mori M (1970) The uncanny Valley. Energy 7:33–35
Mottelson A, Muresan A, Hornbæk K, Makransky G (2023) A systematic review and Meta-analysis of the effectiveness of body ownership illusions in virtual reality. ACM Trans Comput -Hum Interact 30:76:1–7642. https://doi.org/10.1145/3590767
Murray CD, Sixsmith J (1999) The corporeal body in virtual reality. Ethos 27:315–343. https://doi.org/10.1525/eth.1999.27.3.315
Nunes C, Darin T (2024) Echoes of player experience: A literature review on audio assessment and player experience in games. Proc ACM Hum -Comput Interact 8:3041–30427. https://doi.org/10.1145/3677069
Paterson J, Kadel O (2023) Audio for extended realities: A case study informed exposition. Convergence 13548565231169723. https://doi.org/10.1177/13548565231169723
Petkova VI, Khoshnevis M, Ehrsson HH (2011) The perspective matters! Multisensory integration in Ego-Centric reference frames determines Full-Body ownership. Front Psychol 2. https://doi.org/10.3389/fpsyg.2011.00035
Playful Studios (2024) New Super Lucky's Tale [WWW Document]. Playful Studios. URL https://playfulstudios.com/new-super-luckys-tale/ (accessed 3.17.24).
Polyarc Games (2024) Moss [WWW Document]. URL https://www.polyarcgames.com/games/moss (accessed 3.17.24)
Rolla G, Vasconcelos G, Figueiredo NM (2022) Virtual reality, embodiment, and allusion: an Ecological-Enactive approach. Philos Technol 35:1–23. https://doi.org/10.1007/s13347-022-00589-1
Roth D, Latoschik ME (2020) Construction of the virtual embodiment questionnaire (VEQ). IEEE Trans Vis Comput Graph 26:3546–3556. https://doi.org/10.1109/TVCG.2020.3023603
Roth D, Lugrin J-L, Latoschik ME, Huber S (2017) Alpha IVBO - Construction of a Scale to Measure the Illusion of Virtual Body Ownership, in: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '17. Association for Computing Machinery, New York, NY, USA, pp. 2875–2883. https://doi.org/10.1145/3027063.3053272
Sandelowski M (2001) Real qualitative researchers do not count: the use of numbers in qualitative research. Res Nurs Health 24:230–240. https://doi.org/10.1002/nur.1025
Slater M, Banakou D, Beacco A, Gallego J, Macia-Varela F, Oliva R (2022) A separate reality: an update on place illusion and plausibility in virtual reality. Frontiers in Virtual Reality 3.
Southgate E (2020) Conceptualising Embodiment through Virtual Reality for Education, in: 2020 6th International Conference of the Immersive Learning Research Network (iLRN). Presented at the 2020 6th International Conference of the Immersive Learning Research Network (iLRN), pp. 38–45. https://doi.org/10.23919/iLRN47897.2020.9155121
Stitt P, Picinali L, Katz BFG (2019) Auditory accommodation to poorly matched Non-Individual spectral localization cues through active learning. Sci Rep 9:1063. https://doi.org/10.1038/s41598-018-37873-0
Szczotka J, Wierzchoń M (2023) Splitting the unity of bodily self: toward a comprehensive review of phenomenology and psychopathology of heautoscopy. Psychopathology 56:306–314. https://doi.org/10.1159/000526869
The Jamovi Project (2022) Jamovi
Turner A, Berry J, Holliman N (2011) Can the perception of depth in stereoscopic images be influenced by 3D sound? in: Stereoscopic Displays and Applications XXII. Presented at the Stereoscopic Displays and Applications XXII, SPIE, pp. 71–80. https://doi.org/10.1117/12.871960
Ussher M, Cropley M, Playle S, Mohidin R, West R (2009) Effect of isometric exercise and body scanning on cigarette cravings and withdrawal symptoms. Addiction 104:1251–1257. https://doi.org/10.1111/j.1360-0443.2009.02605.x
Vogeley K, May M, Ritzl A, Falkai P, Zilles K, Fink GR (2004) Neural correlates of First-Person perspective as one constituent of human Self-Consciousness. J Cogn Neurosci 16:817–827. https://doi.org/10.1162/089892904970799
Wang X, Mo X, Fan M, Lee L-H, Shi B, Hui P (2024) Reducing Stress and Anxiety in the Metaverse: A Systematic Review of Meditation, Mindfulness and Virtual Reality, in: Proceedings of the Tenth International Symposium of Chinese CHI, Chinese CHI '22. Association for Computing Machinery, New York, NY, USA, pp. 170–180. https://doi.org/10.1145/3565698.3565781
Witmer BG, Singer MJ (1998) Measuring presence in virtual environments: a presence questionnaire. Presence: Teleoperators Virtual Environ 7:225–240. https://doi.org/10.1162/105474698565686
Won AS, Zhou S (2024) Effects of first vs. third-person perspective and self- versus other-avatars on user movements in virtual reality. Int J Hum Comput Stud 187:103259. https://doi.org/10.1016/j.ijhcs.2024.103259
Won AS, Bailenson J, Lee J, Lanier J (2015) Homuncular flexibility in virtual reality. J Computer-Mediated Communication 20:241–259. https://doi.org/10.1111/jcc4.12107
Yee N, Bailenson J (2007) The Proteus effect: the effect of transformed Self-Representation on behavior. Hum Commun Res 33:271–290. https://doi.org/10.1111/j.1468-2958.2007.00299.x
Zhang J, Huang M, Zhao L, Yang R, Liang H-N, Han J, Wang L, Sun W (2020) Influence of Hand Representation Design on Presence and Embodiment in Virtual Environment, in: 2020 13th International Symposium on Computational Intelligence and Design

(ISCID). Presented at the 2020 13th International Symposium on Computational Intelligence and Design (ISCID), pp. 364–367. https://doi.org/10.1109/ISCID51228.2020.00088

Zhang D, Lee EKP, Mak ECW, Ho CY, Wong SYS (2021) Mindfulness-based interventions: an overall review. Br Med Bull Ldab 005. https://doi.org/10.1093/bmb/ldab005