Virtual Perspectives: Effects of Spatial Presence and Agency on Affective and Cognitive Virtual Reality Perspective-Taking

Liyue Da[*]
Department of Computer Science
University of Copenhagen
Copenhagen, Denmark
Department of Computer Science
Aarhus University
Aarhus, Denmark
liyueda@cs.au.dk

Teresa Hirzle
Department of Computer Science
University of Copenhagen
Copenhagen, Denmark
tehi@di.ku.dk

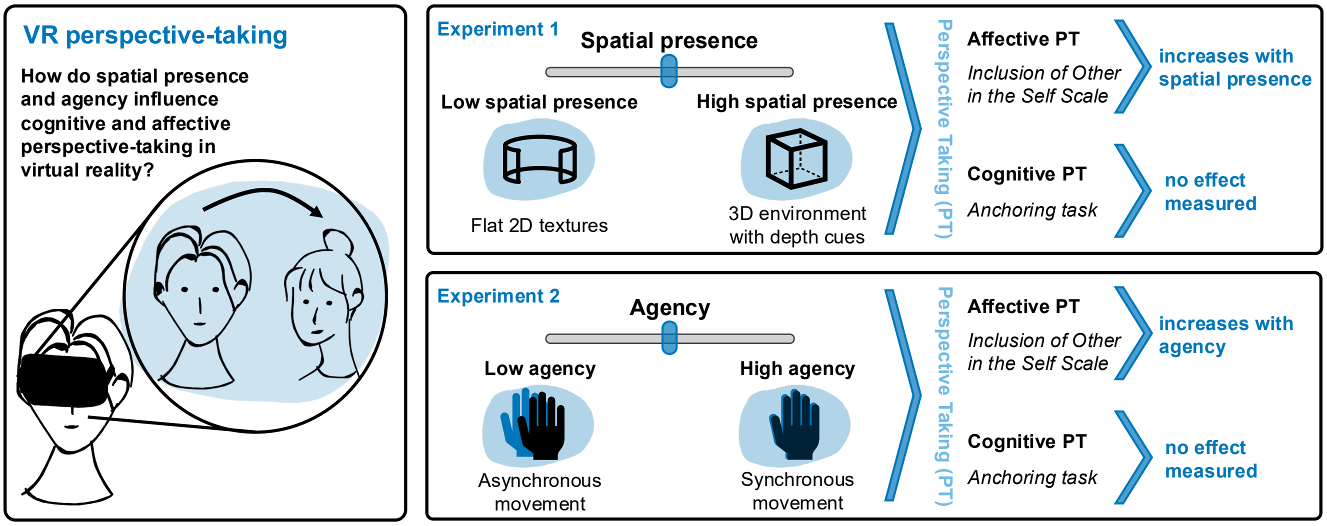


Figure 1: Study overview and main findings. We investigate how spatial presence and agency influence affective and cognitive perspective-taking in a virtual reality perspective-taking study. Users engage in perspective-taking by embodying an avatar and performing a gesture-based task. We independently manipulated spatial presence and agency, two core qualities of VR, to examine their effects on perspective-taking. We found an effect of both spatial presence and agency on affective perspective-taking (i.e., self-avatar overlap), but no effect on cognitive perspective-taking (measured with an anchoring task).

Abstract

Perspective-taking supports empathy, bias reduction, and social cognition, and virtual reality (VR) promises to facilitate it by immersing users into simulated perspectives. Yet, how VR qualities, such as spatial presence and agency, influence – and potentially enhance – perspective-taking remains unclear. We conducted two controlled experiments (N=23 and N=25) using a VR paradigm that manipulated spatial presence (2D vs. 3D) and avatar agency (synchronous vs. delayed motion, and with vs. without a task). We measured cognitive perspective-taking using an anchoring task and self-avatar overlap as an indicator of affective perspective-taking. Results show that increased spatial presence and agency both significantly enhanced affective perspective-taking. However, neither significantly affected cognitive perspective-taking. These findings suggest that while VR qualities can enhance how close a person feels to an avatar, they may not strongly affect cognitive perspective-taking. By studying the influence of two VR qualities, we offer guidance on building more effective perspective-taking experiences in immersive environments.

CCS Concepts

• **Human-centered computing** → **Virtual reality**; **Empirical studies in HCI**; *User studies*; *Social and emotional computing*.

Keywords

Virtual Reality, perspective-taking, spatial presence, agency

---

[*] Corresponding author.

CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/26/04
https://doi.org/10.1145/3772318.3791058

ACM Reference Format:
Liyue Da and Teresa Hirzle. 2026. Virtual Perspectives: Effects of Spatial Presence and Agency on Affective and Cognitive Virtual Reality Perspective-Taking. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 22 pages. https://doi.org/10.1145/3772318.3791058

1 Introduction

Virtual reality is frequently presented as a technology able to fostering empathy and mutual understanding by immersing users in experiences that are rare or inaccessible in their everyday lives (e.g., embodying a child's body [95] or interacting with an age-progressed version of oneself [50]). The ability to adopt another person's point of view is generally referred to as *perspective-taking*. Perspective-taking plays a crucial role in social cognition, empathy, and decision-making [39]. For instance, it enhances conflict resolution by highlighting underlying motivations [24], improves communication by bridging divergent viewpoints [20], and strengthens co-operation by fostering mutual understanding of shared goals [114]. Perspective-taking typically follows either a "self-referential" or "other-referential" approach. The former refers to processing information by linking it to one's own self-concept, using the self as the core reference point for the experience [81]. The latter refers to shifting the reference point to another individual, processing experiences through the lens of that person's perceived characteristics, context, or social identity [94], where social identity describes how a person defines themselves based on group memberships, such as nationality or social class [96]. Virtual Reality (VR) has been suggested as an immersive technology to facilitate perspective-taking since users can literally experience a point of view that is distinct from their own [45, 49]. VR perspective-taking experiences typically also follow either a "self-referential" [41, 103] or "other-referential" approach [1, 49]. In VR, the former involves embodying an avatar that resembles oneself from a different perspective (e.g., aging or future selves [50]); the latter places users in another person's body to experience another person's perspective (e.g., experiencing racial bias [78]). These two techniques have been shown to enhance empathy [45], reduce bias [5], and shift physiological responses or patterns of brain activity related to the self–other boundary [68].

VR perspective-taking studies typically embed the task of *taking a perspective* in highly contextualized narratives, such as embodying a person of a different race [67, 78] or simulating homelessness [70]. Even when the task itself appears neutral, the embodied role often carries social, cultural, or identity-related meanings that participants may draw upon. While such designs can evoke strong emotional and social responses, they also make it difficult to differentiate the effects of the intended manipulation through VR from the situational context. Participants may relate these scenarios to their own social experiences or prior encounters, which introduces uncontrolled variability between individuals. As a result, disentangling the causal effects of specific VR qualities, such as spatial presence and agency, which are often mentioned as motivating reasons for using VR, from the overall impact of VR becomes challenging. Therefore, while several studies have reported that VR perspective-taking scenarios can effectively enhance empathy or reduce bias (e.g., Banakou et al. [5], Herrera et al. [49]), it is largely unknown *why* they work or *what* the exact mechanisms at work are. Questions, such as *Should a person have agency over an avatar they are meant to empathize with? And does the degree of physical presence in the VR environment matter for engaging in perspective-taking?* cannot be answered. Providing systematic findings about how individual VR qualities, such as spatial presence and agency, contribute to perspective-taking, are desirable to be able to design more targeted scenarios, making the experiences more effective and predictable.

We address this gap by systematically manipulating spatial presence and agency as two core VR qualities in a controlled perspective-taking task. This enables us to contribute an analysis of their individual main effects on perspective-taking. This work follows Wirth et al. [108]'s definition of *spatial presence*, who describe it as "a binary experience, during which perceived self-location and, in most cases, perceived action possibilities are connected to a mediated spatial environment, and mental capabilities are bound by the mediated environment instead of reality". This definition differentiates between self-location (SL) and possible actions (PA). Prior work suggests that the feeling of "being there" (the sense of SL) in the virtual environment may be particularly important for perspective-taking, because it could make events feel more psychologically real and personally relevant [58, 89], which might strengthen emotional resonance with virtual others. Therefore, we focus our investigation on the aspect of "being actually there in the environment" of spatial presence. *Agency* refers to "the perception of control over a virtual body" [83]. It is typically discussed within the theoretical model of embodiment by Kilteni et al. [63]. Agency may enhance cognitive and affective perspective-taking, as the ability to intentionally control one's virtual body and actions may increase self-relevance and active engagement [84], evoke motor and predictive processes that underlie understanding others' intentions [57]. Importantly, spatial presence and agency allow users to act in the virtual environment. They thus present a key difference from other technologies (e.g., videos) and could play a key role in VR perspective-taking. Yet, their main effects on VR perspective-taking are unknown.

Summarizing the above, our work is guided by the general research question: *How do spatial presence and agency affect cognitive and affective perspective-taking*? In the evaluation of effects on perspective-taking, we focus on its two key dimensions: *cognitive* and *affective* perspective-taking. *Cognitive perspective-taking* refers to the ability to understand and infer others' thoughts, beliefs, or intentions, emphasizing rational comprehension of "how others think" [18]. *Affective perspective-taking* involves perceiving and sharing others' emotional states, focusing on emotional resonance with "how others feel" [8]. Together, these aspects cover what is generally understood as *psychological perspective-taking*. Since our goal is to provide more fundamental contributions to how VR perspective-taking is formed, we look at the two constructs separately. Although both are often mentioned together as psychological perspective-taking [31], they have different functions and effects [30], so analyzing them separately is valuable for this study. Cognitive perspective-taking is critical because it supports the understanding and prediction of others' thoughts and intentions, forming a cognitive foundation for effective social interaction and communication in virtual contexts. Affective perspective-taking, by contrast, supports emotional connection and empathy, which are key to fostering meaningful engagement with virtual others. Distinguishing between them allows us to clarify how spatial presence and agency might influence different aspects of perspective-taking.

To answer our research question, we conducted two within-participant user studies (N=23 and N=25), investigating the effects of spatial presence and agency on cognitive and affective perspective-taking. Experiment 1 manipulates spatial presence using an immersive 3D environment and a static 2D texture-based environment. In Experiment 2, we manipulate agency by synchronizing and delaying avatar motion. We evaluate cognitive perspective-taking via an anchoring task (reflecting alignment of participants' judgments with the avatar's viewpoint) and affective perspective-taking via self–avatar overlap ratings (capturing perceived emotional overlap). Our findings indicate that both VR qualities are positively associated with affective perspective-taking but show no reliable effect on cognitive perspective-taking, highlighting a dissociation between these components in our context.

In summary, this paper makes two key contributions:

- We provide a systematic investigation of how spatial presence and agency affect cognitive perspective-taking (anchoring strength) and affective perspective-taking (self–avatar overlap), showing that both qualities enhance affective but not cognitive perspective-taking.
- We offer actionable insights on VR perspective-taking: spatial presence and agency both enhance affective perspective-taking, but neither is sufficient for influencing cognitive perspective-taking, which may require explicit guidance or additional cognitive support due to the mental load imposed by immersive environments.

2 Related Work

2.1 Dimensions of Perspective-Taking

"Perspective-taking refers to the ability to recognize another person's point of view [48]." It is fundamental for core social functions, such as inferring intentions, adhering to social norms, and fostering cooperation [71, 99]. It also plays an important role in promoting positive social interactions and strengthening social connections by fostering empathy [7], reducing prejudice [105], and improving cognitive flexibility and social skills [39]. Psychology literature generally dissociates three types of perspective-taking [25, 36]: spatial [72, 102], affective [53, 107], and cognitive [34, 42, 106].

Spatial perspective-taking (sometimes also referred to as visual perspective-taking) is commonly classified into two developmental levels in psychology [33]. Level 1 involves recognizing what others can or cannot see from a given viewpoint [35], and is generally considered to remain egocentric [61]. A canonical example is Piaget's three-mountains task [79], where children are asked to identify what a doll sees from a different position. Level 2 perspective-taking requires mentally representing one's own body relative to another person's position. Individuals simulate how the spatial layout appears from that other person's viewpoint by mentally shifting their own perspective. This level reflects more advanced, allocentric reasoning and is exemplified by mental rotation tasks such as Shepard and Metzler's matching paradigm [87].

Affective, or emotional, perspective-taking (APT), refers to the ability to recognize and interpret others' emotional states based on contextual cues (e.g., a slumped posture, tearful eyes, or trembling voice during a heated argument) [18]. It is strongly associated with empathy, particularly the capacity to respond emotionally to others' distress [64]. Common assessment methods include: narrative-based tasks that ask participants to infer emotions from stories or videos [18]; facial expression recognition tasks [37]; role-playing paradigms where participants adopt another person's role in emotionally charged scenarios [93]; and self-report instruments such as the Empathic Concern subscale of the Interpersonal Reactivity Index (IRI) [64].

Cognitive perspective-taking (CPT) is often situated within the framework of theory of mind (TOM) [80], which refers to the ability to infer others' thoughts, beliefs, feelings, or intentions. This capacity involves interpreting others' internal states and adjusting one's behavior accordingly [52]. A classic assessment task is the *Sally-Anne Task* [6], in which children must predict where an agent will search for an object based on a false belief that contrasts with the child's true knowledge. For instance, in a real-world setting, CPT allows someone to predict a friend's response to a misleading news article by considering their prior beliefs and knowledge, whereas APT helps them recognize the friend's frustration upon discovering the misinformation. Neuroimaging evidence indicates that CPT and APT recruit overlapping yet distinct neural networks, with CPT involving regions such as the medial prefrontal cortex and temporoparietal junction [86].

Conventional tasks for both CPT and APT typically involves reading narratives, watching films, or participating in text-based role-play scenarios [15]. While these methods support inferring others' mental and emotional states, they heavily rely on imagination and prior knowledge, often failing to mitigate egocentric bias. By immersing users in interactive, embodied contexts, VR could reduce the cognitive load of mental simulation.

2.2 Virtual Reality Perspective-Taking

Immersive VR has emerged as a technology for facilitating perspective-taking. Unlike traditional text-based or narrative prompts, VR enables users to act in 3D environments, and thus embody another person's perspective in addition to imagining it. Researchers have leveraged these VR qualities to simulate otherwise inaccessible experiences, such as embodying avatars of different skin colors [44], ages [76], or even species [1]. Prior work indicates that VR has significant effects on enhancing social cognition, empathy, and decision-making. Herrera et al. [49] found that embodying a homeless person in VR elicited more durable attitude change than conventional perspective-taking methods. Didehbani et al. [23] reported that VR-based social cognition training can enhance social cognitive abilities in children with autism. Kandalaft et al. [60] suggested that VR is a promising tool for improving social skills, cognition, and functioning in autism. Borhani and Ortega [12] developed a VR experience for gender swapping and found that it can enhance empathy and reduce stereotype threats in computer science job interviews. Furthermore, a recent meta-analysis indicates that, although the existing evidence is still preliminary, VR shows positive potential in improving learners' reasoning and decision-making performance, particularly in clinical settings [56].

Despite encouraging results, the concrete mechanisms behind *how* perspective-taking in VR works are largely unknown. Most studies target narrowly defined scenarios, and no unified framework links specific VR qualities, such as spatial presence, avatar agency, or locomotion, to distinct dimensions of perspective-taking. Outcome measures also vary widely, lacking consensus on how to assess either cognitive alignment (e.g., decision-making from the avatar's viewpoint) or affective resonance (e.g., emotional connection with the avatar). For example, simulations involving wheelchair navigation increase awareness of mobility barriers [14], but it remains unclear whether such effects generalize to non-locomotion contexts. Similarly, embodiment studies on racial perspective-taking are sensitive to participants' prior attitudes, complicating cross-study comparisons [67].

To investigate these relationships, we systematically manipulate two core VR qualities, *spatial presence* and *agency*, and examine their independent effects on cognitive and affective perspective-taking. Our work focuses on understanding how these VR qualities influence perspective-taking outcomes, rather than on context-specific scenarios. By studying these variables separately, we aim to provide empirical evidence for their individual contributions to VR-based perspective-taking.

2.3 Measuring Cognitive and Affective Perspective-Taking

A variety of methods have been developed in cognitive and social psychology to assess perspective-taking effects. Common approaches include self-report questionnaires, such as the Interpersonal Reactivity Index (IRI) [18], and behavioral paradigms like the Director Task [62] or the dot-perspective task [85]. Questionnaire-based measures are widely used due to their efficiency but may be limited by subjective biases, while behavioral tasks provide more direct evidence of perspective-taking performance, but can be influenced by task-specific demands. Many traditional methods also focus on either affective or spatial aspects of perspective-taking and may not fully capture cognitive perspective-taking.

Building on this foundation, the anchoring paradigm has been proposed as a sensitive, process-oriented measure of cognitive perspective-taking [29]. The anchoring effect, first demonstrated by Tversky and Kahneman [101], occurs when subjective estimates are biased toward an initial anchor value. Epley and Gilovich [28] extended this work by showing that both externally imposed and self-generated anchors can influence judgments. Theoretical accounts such as the anchoring and adjustment heuristic [101] and the selective accessibility theory [75] explain why people insufficiently adjust from anchors. Neuroscience research further indicates that the anchoring effect is associated with activity in brain regions responsible for numerical cognition and executive control [97].

When adapted to perspective-taking, the anchoring paradigm asks participants to estimate values from another's viewpoint, measuring alignment with that perspective's presumed anchor [29]. We use this effect as our primary measure to assess whether VR-based interventions foster greater cognitive alignment with avatars. This implicit, quantifiable method complements traditional measures and offers deeper insight into perspective-taking in immersive environments.

Affective perspective-taking refers to the extent to which an individual emotionally resonates or identifies with another agent [19, 39]. In VR and social cognition research, perceived similarity between the self and another (sometimes also referred to as "self-other overlap") has been proposed as an important component of affective connection, empathy, and the blurring of self–other boundaries. Measuring how similar participants feel to their avatar captures the degree of psychological merging and self-avatar overlap, which are central to affective perspective-taking [19, 39]. Empirical studies have shown that higher reported similarity correlates with greater empathy and affective alignment, and is sensitive to perspective-taking manipulations [4, 31, 47].

3 Experiment 1

Experiment 1 investigated how spatial presence influences cognitive and affective perspective-taking with an avatar in a within-participant design with one independent variable with two levels. The experiment took place in a controlled setting in a quiet room at the University of Copenhagen to ensure consistent hardware and minimize distractions. The experiment had two conditions of each 10 randomized trials, which were presented in randomized order (no block design). We used repeated measures (10 trials per condition) to reduce individual variability and enhance internal validity and statistical power. This approach of employing repeated measures for experimental conditions is commonly done in perspective-taking studies [31].

3.1 Experiment Design

3.1.1 Independent Variable. The experiment had one independent variable (*spatial presence level*) with two levels, operationalized through two different environments. In the *high presence* environment, participants were placed in a realistic 3D-modeled classroom with depth cues (see Figure 2a). Based on spatial presence theory, coherent depth cues (e.g., perspective, occlusion) and familiar semantic content (e.g., desks, chalkboards) enhance the user's sense of "being there" in the virtual space, which is often accompanied by high presence ratings [65, 92]. In the *low presence* environment, participants were placed in an abstract space station with flat 2D textures and minimal depth cues (see Figure 2b). We used a flat, 2D-based design because stripping away depth information and semantic familiarity disrupts the perceptual consistency needed for spatial presence, leading to lower immersion and presence scores [91, 109].

3.1.2 Dependent Variables. We measured three dependent variables: *spatial presence* (to check the effect of the presence manipulation), *cognitive perspective-taking*, and *affective perspective-taking*. Each dependent variable was collected on a per-trial basis, immediately after the embodiment and interaction phase. We used the following measures to assess the dependent variables:

Spatial Presence. We measured spatial presence with a single 7-point Likert item: "I felt like I was actually there in the virtual environment" of the spatial presence experience scale (SPES) [46]. The item is part of the self-location sub scale of the SPES. We chose to use only this item because (1) participants complete 20 trials, so brevity is critical; (2) it is widely validated in VR research with strong reliability and validity [89]; and (3) we did not use interactive

(a) 3D model building scene (high presence condition). **(b) 2D texture scene (low presence condition).**

Figure 2: Experimental environments used to manipulate spatial presence.

environmental manipulations. Therefore, items, such as "It seemed as though I actually took part in the action of the presentation." of the self-location sub scale were not applicable to our experimental setup. It is for the same reason that we considered the possible actions sub scale out of scope for this work.

Cognitive Perspective-Taking (Anchoring Task). We used an anchoring task [101] to measure cognitive perspective-taking. Anchoring tasks have been validated for their use in perspective-taking studies, for example, by Erle and Topolinski [31] and Galinsky and Mussweiler [40]. While embodying the avatar, participants saw a numerical anchor, prompting them to answer an anchoring statement: ("The other person says [anchor]. What do you think?") and provided their own estimate. Anchoring strength was calculated as the standardized z-score across participants and items, reflecting the degree to which responses shifted toward the avatar's perspective.

Affective Perspective-Taking (Self–Avatar Overlap). We measured self–avatar overlap with the Inclusion of Other in the Self (IOS) scale [110]: participants chose one of seven Venn-style diagrams (1 = no overlap, 7 = almost complete overlap), indexing affective perspective-taking via perceived similarity. The IOS scale captures the perceived integration of the avatar into one's self-concept, emphasizing subjective interpersonal closeness rather than objective visual resemblance (e.g., facial features, body shape) [3]. This conceptualization aligns with the core definition of affective perspective-taking, which involves identifying with and understanding another person's emotional experiences [20].

3.1.3 *Hypotheses.*

Based on the theoretical grounding explained in the introduction, we expected that higher spatial presence would lead to higher cognitive and affective perspective-taking. Therefore, we derived the following hypotheses. The detailed rationales for each hypothesis are given below.

H1. Manipulation Check: Participants will report higher *spatial presence* in the high-presence (3D-modeled) environment than in the low-presence (2D-textured) environment.

H2. Participants will exhibit stronger *cognitive perspective-taking* (anchoring strength) in the high-presence environment than in the low-presence environment.

H3. Participants will exhibit stronger *affective perspective-taking* (self–avatar similarity) in the high-presence environment than in the low-presence environment.

Rationales.

H1 Coherent depth cues and semantically rich content enhance the subjective feeling of "being there" so 3D scenes should increase spatial presence [65].

H2 High spatial presence fosters stronger engagement with the virtual avatar's perspective [10]. This deeper engagement may encourage participants to more thoroughly simulate the avatar's judgment process and deliberately adjust their own judgments to align with the avatar's viewpoint, thereby demonstrating stronger cognitive perspective-taking [28].

H3 Stronger spatial presence may focus users' emotional and cognitive resources on the virtual avatar and its perspective, enhancing self-avatar unity [74]. This may make participants more inclined to understand and evaluate virtual events from the avatar's perspective at the emotional and cognitive level, thereby showing higher self-avatar overlap. Compared to traditional non-VR methods, immersive environments that evoke high spatial presence may further strengthen users' identification with the avatar, supporting this expected increase in affective perspective-taking.

3.2 Task and Instructions

To avoid demand characteristics, we withheld the study's true focus until debriefing. The participants were told that the main focus of the experiment is on learning letters from the American Sign Language (ASL) alphabet. Prior work has shown that participants' awareness of experimental hypotheses can substantially bias subjective reports and behavioral responses, including in embodiment and HCI experiments, motivating the use of cover stories as a methodological control [17, 54, 66]. We chose this task to engage the participants in a body-based task (i.e., seeing their hands while learning the gestures).

Participants received the following instruction via an in-VR panel at the start of each session:

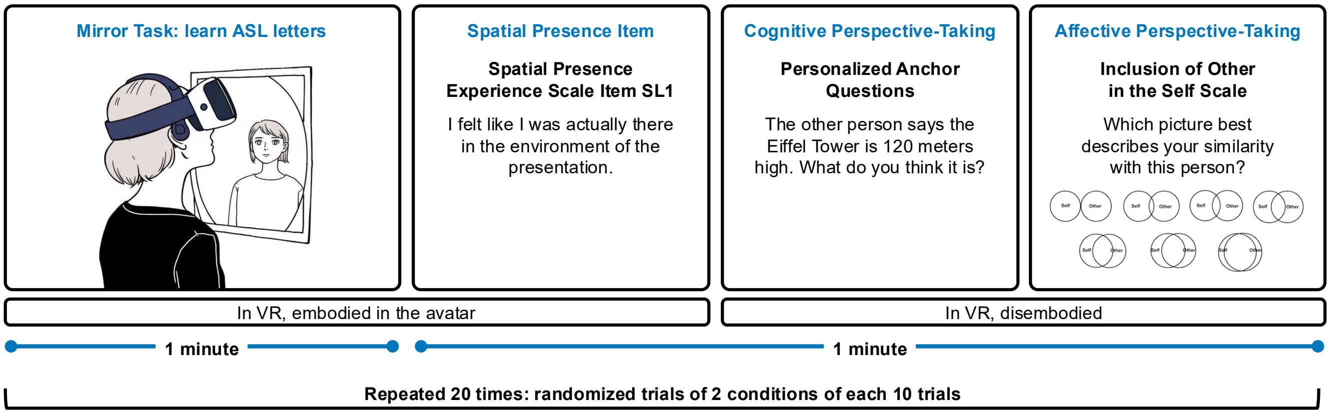


Figure 3: Experiment procedure: each of the 20 trials includes four phases: (1) embodiment and ASL task, (2) presence rating, (3) cognitive perspective-taking task (anchoring task), and (4) affective perspective-taking task (self-avatar overlap score). The trials were presented in a randomized order. The figure uses an image generated with DeepAI[1].

> *"In this experience, you will help another person learn letters from the American Sign Language (ASL) alphabet. To do this, you will be embodied in the avatar you are seeing. This means that your movements will be aligned with the avatar's movements."*

The experimenter then restated these instructions. After the experiment, several participants stated that they thought the main task of the study was on learning the ASL alphabet. None of them discovered the true purpose of the study.

3.3 Procedure

Participants first gave their informed consent. Then, they completed a demographics questionnaire, which measured the following aspects: participants' gender, age, history of neurodivergence (e.g., ADHD, Autism etc.), and previous experience with VR. Participants unfamiliar with VR completed two practice trials using an avatar that was not used in the study to learn the controls. Figure 3 shows the procedure in detail. Each trial consisted of the following phases:

(1) **Embodying the avatar:** Participants embodied an avatar and were tasked to learn ASL letters, as shown in the instructions. They received visual instructions of how to perform the gesture for each ASL letter through a 2D drawing of the gesture shown in front of them on a screen in the VR environment. During the embodiment phase, participants saw the avatar they embodied in a mirror in front of them. This phase took one minute and participants learned letters from the standard 26-letter ASL alphabet. ASL letters were randomly chosen from the ASL alphabet. We excluded six letters (H, K, M, N, T, and R), because they involve significant overlap of fingers, which was problematic with the headset's limited hand tracking and led to problems in displaying the hands correctly.

(2) **Spatial presence rating:** After the embodiment phase, participants answered the spatial presence question while embodying the avatar. We did this to keep their feeling of presence tied to the experience without interrupting it by disembodying the avatar (Figure 4a).

(3) **Disembodying the avatar:** Next, participants disembodied the avatar and answered the remaining questions. This phase was included to ensure that participants clearly distinguished between themselves and the avatar to separate the responses on the perspective-taking outcome from the task.

(4) **Cognitive perspective-taking (anchoring task):** Then, participants indicated their level of cognitive perspective-taking by doing the anchoring task. They were asked to respond to the anchoring questions that were formulated as follows: *"The other person says [anchor]. What do you think?"*

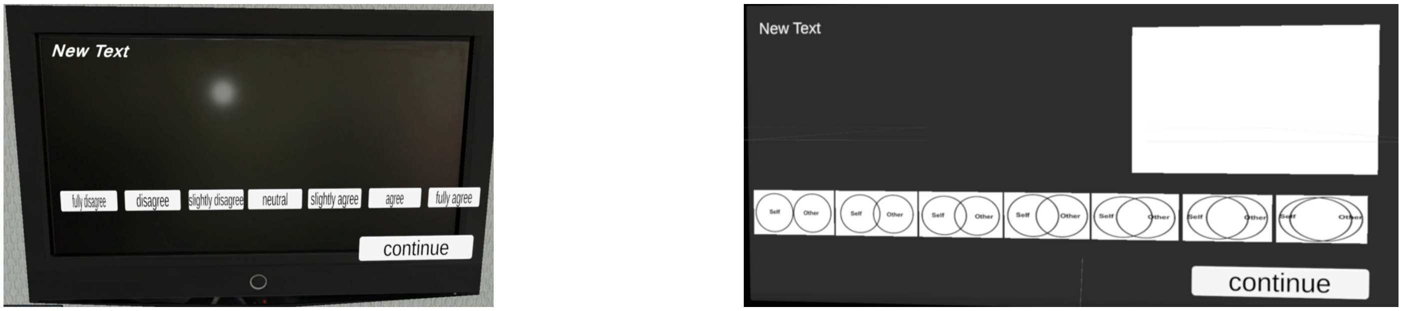
(5) **Affective perspective-taking (self-avatar overlap):** The final activity in each trial was to indicate the level of affective perspective-taking. Participants indicated how similar they felt to the avatar through the IOS scale (Figure 4b).

3.4 Validity and Controlling for Biases

We fully randomized all trials across conditions within each participant to mitigate order and learning effects [32, 43]. We minimized experimenter expectancy bias [82] by delivering identical, neutral instructions ("The other person says [anchor]. What do you think?") across all trials. All anchor questions followed established anchor design principles, using questions with vague but potentially familiar facts [28], designed to evoke vague intuitions rather than explicit knowledge [75], and worded in a neutral way [13]. Furthermore, we increased statistical power and lowered the risk of Type II errors by running 10 trials per condition [16]. To prevent familiarity effects from repeated exposure to a single face [98], we randomly assigned one of 20 visually similar avatars in each trial (see Figure 5). This variability not only reduces recognition and preference biases but also enhances ecological validity, aligning with Slater and Steed's recommendation that simulations reflect realistic human diversity [91].

---

[1] The person wearing a VR headset and the person in the mirror in the left panel of the procedure figure were generated with the DeepAI image generator https://deepai.org/machine-learning-model/text2img.



(a) The presence question interface.

(b) The post-embodiment questionnaire interface.

Figure 4: Questionnaire interface. The placeholder fields labeled "New Text" are Unity defaults that were dynamically replaced with the actual question text during task execution. We show the static interface here for illustration, while the real content was only visible during runtime.

3.5 Implementation

We developed a standalone VR application in Unity for the Meta Quest 2, featuring 20 scenes (10 fully modeled 3D and 10 textured 2D) presented in randomized order. Participants were represented by realistic humanoid avatars. The avatars were imported from the Unity Asset Store, customized in Blender (hair mesh and clothing color), and retargeted with the Meta Movement SDK to enable upper-body tracking without external devices. The mirror-based learning task displayed English letters and corresponding hand gestures on a virtual TV above the participant. One letter was shown for 10 seconds; a secondary Unity camera rendered the user's movements to a Render Texture on a plane to simulate a mirror. Built-in hand tracking supported natural pointing and selection gestures via Unity's event system. After each trial, in-VR questionnaires captured subjective responses, which were automatically logged to CSV with trial-level metadata (e.g., avatar ID, presence condition); no continuous behavioral data was recorded.

Figure 5: Avatar set and usage across experiments. The red box marks the 20 avatars used in Experiment 1; Experiment 2 used the full set (24). In each trial, participants embodied a different avatar, and avatar assignment to trials was randomized.

3.6 Results

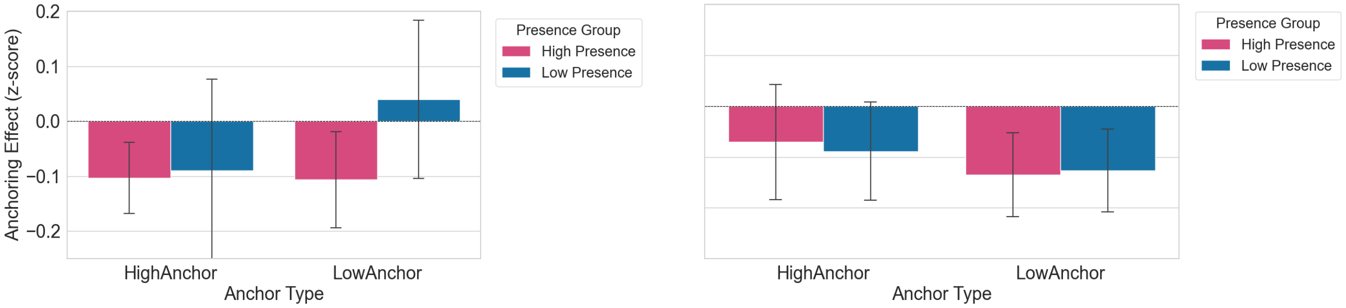
3.6.1 Participants. 23 participants (14 female, 9 male, age range 23–32) completed the study. All participants were recruited via university mailing lists and received a small compensation for their participation. An a-priori power analysis for a paired-samples t test using G*Power[2] (Cohen's $d_z = 0.50$, one-tailed $\alpha = 0.05$, $1 - \beta = 0.8$) yielded a sample size of 27. Due to resource constraints, the final sample of experiment 1 was slightly smaller, resulting in reduced statistical power. This limitation was accepted because both experiments were essential to addressing the overall research question, requiring us to allocate the available resource across both studies. Participants were required to have normal or corrected-to-normal vision and no known neurological or motor impairments. Before the experiment, each participant completed a demographic questionnaire indicating age, gender, prior VR experience, and any relevant medical conditions, and sign an informed consent form.

3.6.2 Data Analysis. We conducted a two-step analysis to assess the effect of presence on cognitive perspective-taking and affective perspective-taking. First, we conducted a manipulation check, comparing spatial presence ratings between the 3D-modeled and 2D-textured environments. Then, we examined whether this perceived presence influenced (1) cognitive perspective-taking and (2) affective perspective-taking. To compute anchoring effects, both the participants' estimates and the anchor values were z-scored within each question–condition group. Specifically, for each question type (e.g., age estimation, salary estimation), we calculated the mean and standard deviation across all participants and conditions. Each participant's estimate and the corresponding anchor value were then standardized using these group-level statistics: $z = (x - \mu)/\sigma$, where x is the raw value, μ is the group mean, and σ is the group standard deviation. The anchoring effect was then calculated as the difference between the standardized estimate and the corresponding standardized anchor.

To ensure high data quality, we excluded 9 trials with response z-scores $|z| > 3$, indicating extreme outliers. We also dropped the 25% of anchor questions with the lowest coefficient of variation (5

---

[2] G*Power: https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower



(a) Anchoring Effect by Anchor Type and Presence Group

(b) Anchoring Effect by Anchor Type and Experimental Condition

Figure 6: Anchoring effects (cognitive perspective-taking) under different grouping strategies.

items), due to limited variability that would not provide meaningful anchoring effects.

3.6.3 Summary of the Results. We summarize the main results here and present the details about the statistical tests in the following sub sections. Our findings confirm that 3D modeled scenes significantly increased subjective presence ratings compared to 2D textured scenes. This increase in presence was associated with significantly higher affective perspective-taking ratings (perceived overlap with the avatar): trials labeled as *HighPresence* (based on median-split self-reported presence ratings) had significantly higher overlap scores than *LowPresence* trials, suggesting an effect of spatial presence on affective perspective-taking.

In contrast, the increase in presence did not significantly influence participants' cognitive perspective-taking ratings as measured by anchoring effects. Taken together, while increases in spatial presence caused higher affective perspective-taking scores , we did not find an effect of spatial presence on cognitive perspective-taking.

3.6.4 Spatial Presence Manipulation. Participants reported significantly higher presence in the 3D condition ($M = 5.25$, $SD = 1.28$) compared to the 2D condition ($M = 4.84$, $SD = 1.45$), confirmed by a paired t-test ($t = 2.87$, $p = 0.0089, \eta^2 = 0.27$). Data were approximately normally distributed according to a Shapiro-Wilk test ($W = 0.938$, $p = 0.1616$).

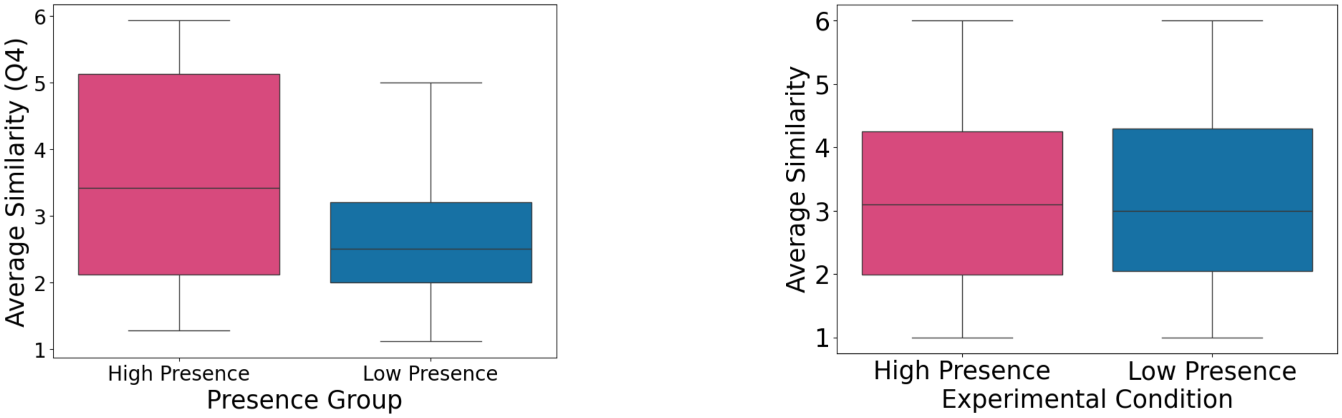
3.6.5 Cognitive Perspective-Taking. We verified that our dependent measure – the standardized difference between each judgment and the corresponding anchor – was approximately normally distributed across conditions (Shapiro–Wilk tests all $p > .05$). We then proceeded with parametric analyses. To examine a potential effect of spatial presence on cognitive perspective-taking, measured through anchoring strength, we grouped the data in two ways. First, we grouped the presence ratings by condition (2D vs. 3D scene). Secondly, we grouped them by the median split of participants' reported presence scores. The second grouping ensured that we capture individual influences, since presence has been shown to be a highly subjective experience. For both grouping strategies, we employed two parametric approaches appropriate for a within-participant design and normally distributed data: (1) Paired-samples t-tests to compare anchoring effects across presence levels at the participant level and (2) Linear mixed-effects models with subject-level random intercepts, which accounted for repeated measures and trial-level variation in anchor values.

The paired-samples t-test based on the condition-based grouping did not reveal a statistically significant result ($t = 0.03$, $p = 0.98$) with means of: High/HighAnchor $M = -0.071$, High/LowAnchor $M = -0.135$, Low/HighAnchor $M = -0.088$, and Low/LowAnchor $M = -0.127$ (Figure 6b). Similarly, the grouping based on a median split did also not reveal a statistically significant difference in anchoring effects ($t = -1.10$, $p = 0.29$). The corresponding means are: HighPresence/HighAnchor $M = -0.103$, HighPresence/LowAnchor $M = -0.106$, LowPresence/HighAnchor $M = -0.090$, and LowPresence/LowAnchor $M = 0.040$ (Figure 6a). Moreover, the multiple linear regression analyses in both grouping approaches did not reveal any significant main or interaction effects (all $p > 0.05$; see Table 1 for full results).

In summary, regardless of the grouping method used, both the paired-samples t-tests and the mixed linear models did not detect any significant main or interaction effects. This indicates that, in the present study, the expected level of presence did not play the anticipated role in modulating the standardized difference between the participant's judgment and the anchor value.

We further explored how presence ratings and anchor direction influenced cognitive perspective-taking (measured with anchor strength) with multiple linear regression and a two-way ANOVA. Rejecting our hypothesis that higher presence would increase participants' adjustment toward the anchor, neither the regression nor the ANOVA revealed significant main or interaction effects ($p > 0.05$ for all tests). This aligns with our earlier paired-samples t-test and mixed-effects model analyses, all failing to detect systematic differences. In sum, the expected level of presence did not modulate the standardized anchoring effect in this study. Details about the tests are shown in the Appendix in subsubsection A.3.1.

3.6.6 Affective Perspective-Taking. We measured affective perspective-taking by assessing perceived self-avatar overlap with the IOS scale. We examined potential effects under two grouping approaches. Similarly to the analysis of the presence scores, in the first approach, we compared overlap scores across the 2D and 3D scene conditions, treating *Scene Type* (2D vs. 3D) as a within-participant factor. A paired-samples t-test revealed no significant difference ($t = 0.70$, $p = 0.49$) with mean scores of $M_{\text{2D}} = 3.26$ ($SD = 1.52$) and



(a) Self-avatar overlap scores by presence group **(b) Self-avatar overlap scores by experimental condition**

Figure 7: Comparison of average overlap scores (affective perspective-taking) under two grouping strategies: (a) based on participants' reported presence (median split), and (b) based on experimental condition (2D vs. 3D).

$M_{\text{3D}} = 3.20$ ($SD = 1.56$) (Figure 7b). For the second grouping, we split the trials into high- and low-presence conditions based on the median-split of each participant's presence ratings. As each participant contributed data to both high-presence and low-presence conditions, we used within-group tests. A paired-samples t-test showed a significant difference in overlap scores between the high- and the low-presence group ($t = 4.68$, $p = 0.0002$) with mean overlap scores of $M_{\text{HighPresence}} = 3.61$ ($SD = 1.60$) and $M_{\text{LowPresence}} = 2.72$ ($SD = 1.18$) (Figure 7a).

3.7 Discussion

Several factors may explain why the observed increase in presence did not significantly affect participants' cognitive perspective-taking.

3.7.1 Effect Size and Threshold Effects. Participants reported significantly higher presence in the 3D modeled scene compared to the 2D texture scene, possibly due to the 3D scene's higher ecological validity. Although this difference was statistically significant with a medium-to-large effect size ($\eta^2 = .27$), the absolute difference in mean presence ratings was modest (3D: $M = 5.25$, $SD = 1.28$; 2D: $M = 4.84$, $SD = 1.45$). This suggests a potential threshold effect: while the change in presence was meaningful statistically, its modest magnitude in raw scores may be insufficient to trigger notable psychological or behavioral chain reactions. In other words, the limited enhancement in perceived presence may not be enough to surpass critical thresholds in cognitive processing or self-perception, thereby failing to significantly promote cognitive perspective-taking.

3.7.2 Cognitive Resource Allocation Conflict. High immersion may demand additional attentional and cognitive resources for perceptual processing, leaving fewer available for effortful numerical adjustment of the anchor. According to the anchoring–adjustment framework, such adjustment relies on top-down control and working memory [27]. Immersive VR can also increase cognitive load compared to desktop simulations [69], which suggests that under conditions of high presence, participants could have insufficient "bandwidth" to carry out careful adjustments. As a result, their estimates might deviate more from the anchor, leading to larger absolute errors.

Taken together, these factors suggest that while presence enhances affective engagement with avatars, its influence on cognitive perspective-taking may depend on stronger manipulations, more sensitive tasks, and lower competing cognitive demands.

4 Experiment 2

Experiment 2 investigated how agency affects cognitive and affective perspective-taking in VR using the same mirror task and a repeated-measures design. It manipulated two within-participant variables: *agency level* (synchronous vs. delayed control) and *task presence* (with vs. without a sign language learning task). By including the task manipulation, we aimed to examine whether increased cognitive load or goal-directed activity would modulate participants' sensitivity to agency disruptions based on recent findings suggesting that task engagement may affect body ownership [73], and thus indirectly agency. This factorial design aimed to clarify how agency interacts with cognitive engagement during avatar interaction.

4.1 Experiment Design

4.1.1 Independent Variables. The experiment has two independent variables, both manipulated within-participant: agency level and task condition.

Agency Level: This variable was operationalized through real-time synchronization of avatar movements with participant actions. In the *high agency* condition, the avatar's movements were synchronized with the participant's actions in real time, preserving real-time spatial and temporal alignment. In the *low agency* condition, a 2-second delay was introduced to the avatar's joints to induce spatial and temporal misalignment, disrupting agency.

To manipulate agency, we referred to Iarygina et al. [55]'s review of VR techniques that reduce agency: (1) *prerecorded/external*,

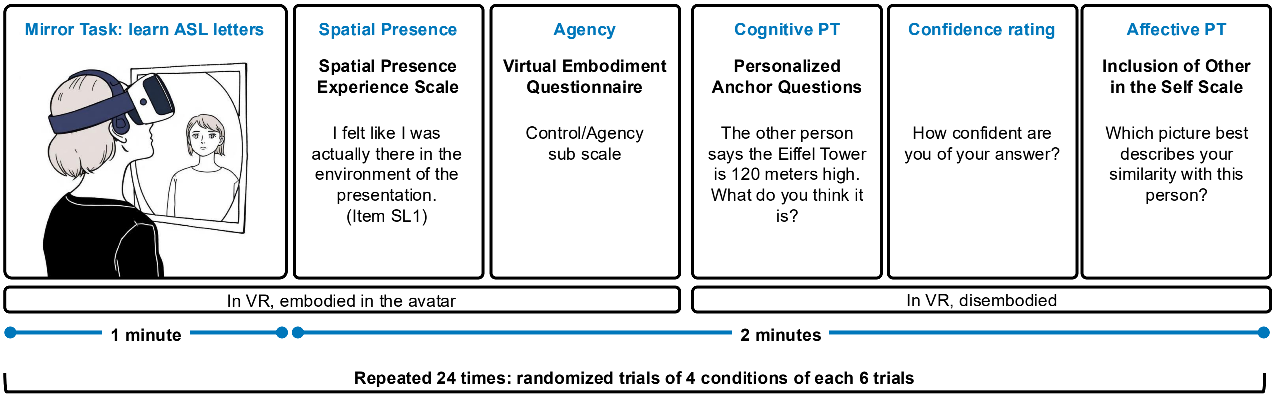


Figure 8: Experiment 2 procedure: each of the 24 trials includes the five phases (1) presence rating, (2) agency assessment, (3) anchoring task, (4) confidence rating, and (5) overlap score. The figure uses an image generated with DeepAI[1].

(2) *delayed*, and (3) *dislocated* movements. We excluded the first and third methods for both technical and conceptual reasons. The dislocated approach would require spatial offsets in hand or finger positions that, if too large, would result in anatomically implausible postures, and if too small, might be perceived by participants as tracking errors rather than intentional manipulations. The prerecorded/external condition was also unsuitable because it conflicted with the task of learning ASL letters. Completely replacing their movements with prerecorded actions would make this task implausible and could draw attention to the manipulation, thereby increasing demand characteristics.

Task Condition: This variable controlled the cognitive load and movement constraints during the embodiment experience. In the *task present* condition, a TV screen displayed American Sign Language (ASL) hand gestures alongside corresponding images, and participants imitated these signs virtually. In the *no task* condition, the TV screen remained blank, and participants freely performed spontaneous movements. This manipulation served two purposes: (1) performing a goal-directed task increased cognitive load, potentially diverting attention from avatar responsiveness; (2) in the no-task condition, unconstrained movements may have enhanced the perceived self-relatedness of the avatar, increasing agency and embodiment.

4.1.2 Dependent Variables. We measured five dependent variables: *spatial presence* (manipulation check), *agency* (manipulation check), *cognitive perspective-taking*, *confidence rating*, and *affective perspective-taking*. Each dependent variable was collected on a per-trial basis, immediately after the embodiment and interaction phase.

As in Experiment 1, spatial presence was assessed with a single item from the Spatial Presence Experience Scale [46]; cognitive perspective-taking was measured with the personalized anchoring task with standardized z-scored anchoring strength; and affective perspective-taking was assessed using the IOS scale [110].

Agency. Perceived agency was measured with four items (CO1–CO4) from the Control/Agency subscale of the Virtual Embodiment Questionnaire (VEQ) [83]. Participants rated their agreement on a 7-point Likert scale (1 = fully disagree, 7 = fully agree) with statements such as "I felt like I was controlling the movements of the virtual body." A composite agency score was computed by averaging the four items, with higher values indicating a stronger sense of control over the avatar.

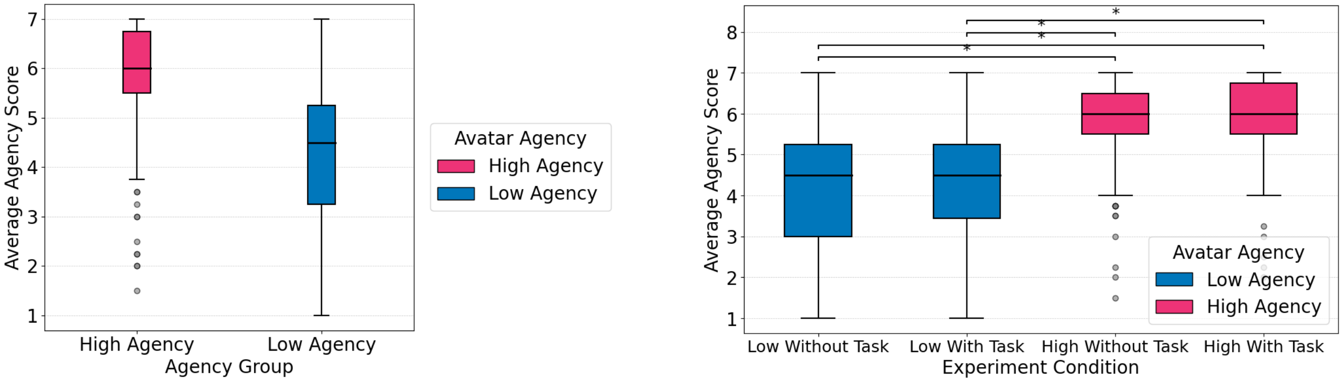
Confidence Rating. Immediately after providing each estimate in the anchoring task, participants rated their confidence on a 7-point Likert scale (1 = not at all confident, 7 = very confident). This per-trial confidence score served as an auxiliary index for cognitive perspective-taking in subsequent analyses.

4.1.3 Hypotheses. Based on the theoretical grounding explained in the introduction, we expected that higher agency would lead to stronger cognitive and affective perspective-taking. Therefore, we derived the following hypotheses. The detailed rationales for each hypothesis are given below.

H1. Manipulation Check: Participants will report lower *agency* in the misaligned (low-agency) condition than in the high-agency condition.
H2. Participants will exhibit stronger *cognitive perspective-taking* (anchoring strength) in the high-agency condition than in the low-agency condition.
H3. Participants will exhibit stronger *affective perspective-taking* (self–avatar overlap) in the high-agency condition than in the low-agency condition.

Rationales.

H1 Delays disrupt sensorimotor predictions [38]; a 2-second delay was expected to significantly reduce perceived agency.
H2 A reduced sense of agency may hinder cognitive alignment with the avatar. When control over the avatar is less immediate or less fluent, participants may be less likely to fully adopt the avatar's viewpoint when making judgments. Prior work suggests that agency enhances embodiment and facilitates anchoring to the avatar's perspective [4].
H3 Affective perspective-taking involves affective identification and empathic understanding of another agent. While direct



(a) Boxplots showing participants' reported sense of agency under *Low* and *High Agency* conditions.

(b) Boxplot of agency scores across the four conditions, with significant pairwise comparisons indicated by asterisks.

Figure 9: Boxplots illustrating self-avatar overlap scores and agency scores across different experimental conditions.

empirical evidence linking agency to affective perspective-taking is limited, previous work suggests that agency contributes to body ownership and self-avatar merging – processes that underlie social bonding and empathic alignment. For example, You et al. [113] found that participants' control over their avatar influenced their bodily and cognitive responses in VR. Reduced agency may weaken emotional alignment, decreasing affective perspective-taking.

4.1.4 Task and Instructions. Participants received the following instruction via an in-VR panel at the start of each session:

> *"In this experience, you will help another person learn letters from the American Sign Language (ASL) alphabet. To do this, you will be embodied in the avatar you are seeing. This means that your movements will be aligned with the avatar's movements."*

The experimenter then restated these instructions. To avoid demand characteristics, we withheld the study's true focus until debriefing.

4.1.5 Procedure. The procedure of Experiment 2 followed the same structure as Experiment 1, with the following modifications (Figure 8):

1. **Agency Assessment:** After completing the presence questions, participants responded to four agency questions (CO1–CO4) from the VEQ while remaining in the embodiment state.
2. **Confidence Rating:** Following the anchoring task, participants rated their confidence in their answer on a 7-point Likert scale.

Figure 5 shows the avatars used in Experiment 2. Avatars were shown in first-person during embodiment and third-person during the anchoring task.

4.1.6 Improving Response Validity. In Experiment 1, we did not prevent participants from copying the avatar's answer, which may have encouraged participants to follow the avatar's assessment rather than their own.

To address this in Experiment 2, we made two changes: (1) participants were explicitly told **not to replicate the anchor**; and (2) they rated their **confidence** (1–7 Likert scale) after each estimate, allowing identification of passive responses.

4.1.7 Implementation. This experiment used the same system as Experiment 1, with the only change being the addition of a delayed movement condition. The VR application was developed in Unity for Meta Quest 2. All 24 3D-modeled scenes were presented in random order, and responses were logged in `.csv` files. To reduce agency, we introduced a 2-second delay between user movement and avatar response. For each upper-body joint, we implemented a movement buffer that stored position and rotation at each frame, rendering joint states from several frames earlier to achieve the delay. All other components, including avatar appearance, mirror rendering, hand tracking, ASL mirroring task, and questionnaires, remained unchanged from Experiment 1.

4.2 Results

4.2.1 Participants. 25 participants (11 male, 14 female; age range 23–39) completed the study. All participants were recruited through university mailing lists and received a small compensation for their participation. An a-priori power analysis using G*Power[2] (effect size $f = 0.25$, $\alpha = .05$, power = .80) for a repeated-measures ANOVA (1 group, 4 measurements, correlation among repeated measures = .50) indicated that a minimum of 24 participants would be sufficient to detect medium-sized effects. We enrolled 25 participants; however, one participant always selected "fully disagree" for every questionnaire item irrespective of content and was excluded based on a data-quality criterion, yielding a sample of $n = 24$ participants included in the analysis. With $n = 24$, planned power for detecting $f = .25$ effects remains approximately .80.

Participants were required to have normal or corrected-to-normal vision and no known neurological or motor impairments. Before the experiment, each participant completed a demographic questionnaire indicating age, gender, prior VR experience, and any relevant medical conditions, and sign an informed consent form.

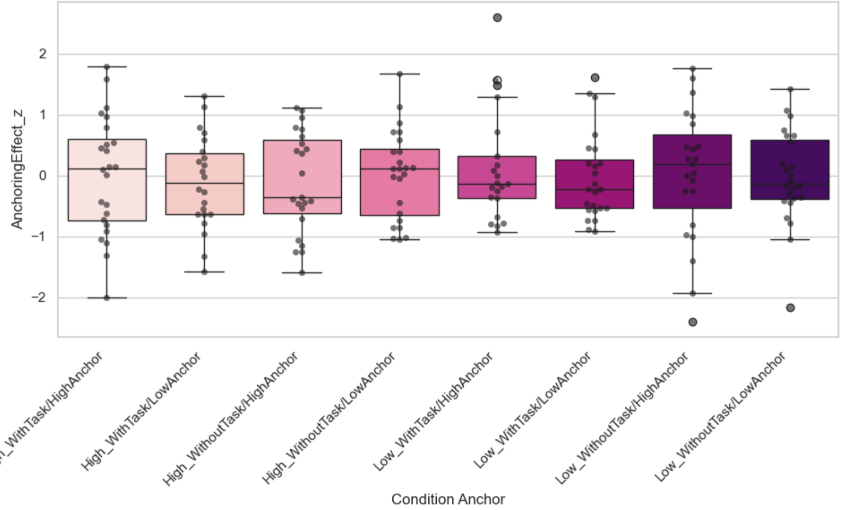


Figure 10: Distribution of the confidence-weighted *AnchoringEffect_z* across the eight Condition_Anchor combinations, where High and Low refer to the level of *avatar agency.*

4.2.2 Data Analysis. Similar to Experiment 1, we conducted a two-step analysis. First, we conducted a manipulation check comparing participants' agency ratings under simultaneous and 2-s delay conditions. Then, we examined whether this affected (1) cognitive perspective-taking and (2) affective perspective-taking.

To assess whether avatar agency and task presence influenced anchoring effects, we extended the calculation from Experiment 1 by incorporating participants' confidence ratings. Specifically, we computed a confidence-weighted standardized anchoring score for each trial by multiplying the z-scored anchoring value by the participant's self-reported confidence (scaled between 0 and 1).

No trials were excluded as extreme outliers ($|z| > 3$). We also dropped the 25% of anchor questions with the lowest coefficient of variation (6 items), due to limited variability that would not provide meaningful anchoring effects.

4.2.3 Summary of the Result. We summarize the main results here and provide more details on the statistical tests in the following sub sections. Our findings confirm that the agency manipulation was successful: participants reported significantly higher agency in the *High* compared to the *Low-agency* conditions. Both *High-agency* conditions showed significantly higher agency ratings than both *Low-agency* conditions, whereas task presence did not affect agency ratings.

In terms of affective perspective-taking, self-avatar overlap scores were significantly higher in *High-agency* conditions than in *Low-agency* conditions. We also found interaction effects of task presence and agency on affective perspective-taking. Participants indicated higher affective perspective-taking ratings in the *High-agency* condition with a task compared to both low agency conditions (with and without task). Furthermore, the *High-agency* condition in which a task was present yielded higher ratings compared to the *Low-agency* condition with a task present.

In contrast, agency and task manipulations did not significantly influence cognitive perspective-taking. The anchoring effect remained near zero across all conditions, and no differences emerged between HighAnchor and LowAnchor items.

Taken together, while higher agency significantly increased affective perspective-taking (measured with IOS scale), it did not affect cognitive perspective-taking, measured by anchoring effects.

4.2.4 Agency Manipulation. Agency scores (averaged across four agency questions) did not follow a normal distribution (Shapiro–Wilk and D'Agostino's tests, both $p < .001$), so we conducted nonparametric analyses. Wilcoxon signed-rank tests revealed significantly higher agency in *High* ($M = 5.88, SD = 1.02$) compared to *Low* conditions ($M = 4.25, SD = 1.43$, $W = 1494.5$, $p < .001$) with a large effect size ($r = 0.79$). In the synchronous control condition, most scores were near the upper limit of the scale, and a few outliers did not change this overall pattern.

To examine whether the presence of a task influenced agency scores across the four conditions, we conducted a Friedman test, which yielded $\chi^2(3) = 48.50$, $p < 0.001$, indicating a significant overall effect. Bonferroni-corrected Wilcoxon signed-rank tests showed that both High-agency conditions were rated significantly higher than both Low-agency conditions (all $p_{\text{adj}} < .001$; mean differences ≈ 1.35–1.58), while adding a task did not change agency ratings within the same agency level (Low: $p > 0.05$; High: $p > 0.05$). Full statistics are provided in Appendix Table 3; the corresponding boxplot is shown in Figure 9b.

4.2.5 Cognitive Perspective-Taking. The paired differences in standardized anchoring effects between the High and Low agency conditions followed a normal distribution (Shapiro–Wilk test: $W = 0.949, p = 0.261$; D'Agostino's K^2 test: $K^2 = 1.879, p = 0.391$), so we used a repeated-measures ANOVA for the analysis.

We conducted a 2×2 repeated-measures ANOVA with within-participant factors *Agency* (High vs Low) and *Task* (WithTask vs

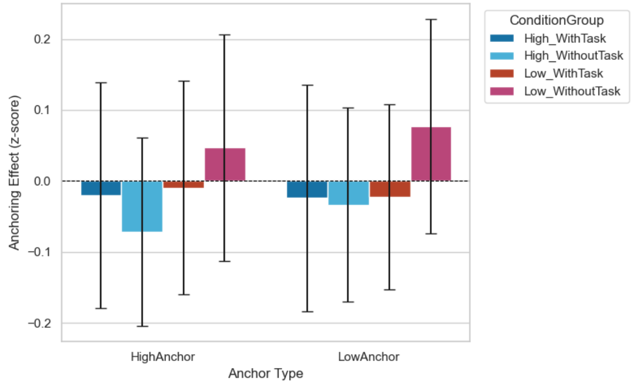


Figure 11: Mean confidence-weighted *AnchoringEffect_z* (±SE) by AnchorType and task condition, where High and Low refer to the level of *avatar agency.*

WithoutTask) on the confidence-weighted standardized anchoring effect. The ANOVA revealed no significant main effect of *Agency* ($F(1, 23) = 0.0312$, $p = 0.86$), no significant main effect of *Task*, ($F(1, 23) = 0.0516$, $p = 0.82$), and no significant interaction effect ($F(1, 23) = 0.0505$, $p = 0.82$). Full ANOVA statistics are provided in Appendix Table 4.

To further determine whether anchor type (HighAnchor vs. LowAnchor) influenced the weighted anchoring effect, we analyzed each participants' mean confidence-weighted *AnchoringEffect_z* separately for high-anchor and low-anchor items within each task condition. Figure 10 displays the distribution across the eight Condition × Anchor combinations, and Figure 11 shows the cell means with standard error bars.

The confidence-weighted anchoring effect was near zero in all conditions (-0.078 to 0.047; see Table 5). Bonferroni-corrected paired-samples t-tests revealed no significant differences between any HighAnchor and LowAnchor comparisons (min $p = 0.4183$, all $p > 0.42$), indicating that neither anchor type nor task condition systematically affected the confidence-weighted anchoring effect.

4.2.6 *Affective Perspective-Taking.*

The IOS scores significantly deviated from a normal distribution (Shapiro–Wilk test: $W = 0.9342$, $p < 0.001$); D'Agostino's K^2 test: $K^2 = 183.43$, $p < 0.001$). Accordingly, we proceeded with nonparametric analyses.

To examine the effects of agency, task, and their interaction on affective perspective-taking, we conducted a nonparametric factorial repeated-measures ANOVA using the Aligned Rank Transform (ART) procedure. This analysis revealed significant main effects of *Agency* ($F(1, 23) = 5.03$, $p = .035$) and *Task* ($F(1, 23) = 7.07$, $p = .014$), as well as a significant interaction between the two ($F(1, 23) = 35.25$, $p < .001$). The results indicate that IOS scores were significantly higher in the *High Agency* compared to *Low Agency* conditions. Furthermore, the scores were higher in the *With Task* compared to *Without Task* conditions. Finally, to examine interaction effects, follow-up pairwise Wilcoxon signed-rank tests (FDR-corrected) showed that self-avatar overlap scores in the *High Without Task* condition were significantly higher than in both the *Low With Task* ($W = 17.50, p = .0039$) and *Low Without Task* ($W = 30.50, p = .0188$) conditions (see Figure 12). In addition, overlap scores in the *High With Task* condition were significantly higher than in the *Low With Task* condition ($W = 35.00, p = .0188$). Full results of the ART ANOVA and descriptive statistics are presented in the Appendix in Table 7 and Table 6.

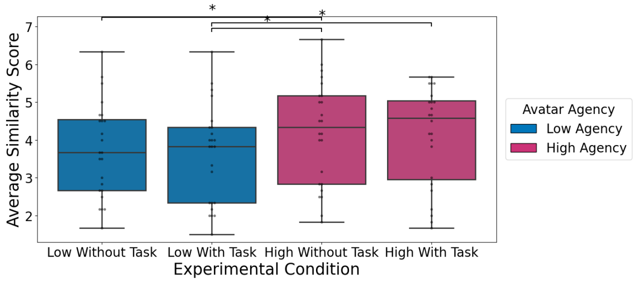


Figure 12: Boxplots showing average affective perspective-taking scores for each experimental condition. Significant pairwise differences (FDR-corrected, $p < .05$) are indicated by asterisks. Updated to more accessible palettes.

4.3 Discussion

4.3.1 *Agency and Self–Avatar Overlap.*

Our findings demonstrate that agency significantly influences participants' perceived self-avatar overlap. Higher agency consistently increased affective perspective-taking, suggesting subjective control enhances self-avatar merging. This aligns with prior work indicating that agency is a key component of embodiment [59, 100], and that active control leads to stronger self-identification with digital bodies.

Interestingly, while both *High Agency* conditions were rated similarly, the *High Without Task* condition produced the highest self-avatar overlap scores overall, even surpassing *High With Task*. This suggests that although task engagement is often thought to increase embodiment [73], in our case, removing the task may have allowed participants to focus more on the avatar and the sensation of agency itself, uninfluenced by external goals. The finding suggests further exploration into how task relevance interacts with agency to affect self-avatar merging.

4.3.2 *Agency-Induced Autonomy May Suppress Susceptibility to Social Influence.*

We speculate that the lack of significant anchoring effects in the *High Agency* condition may reflect a heightened sense of autonomy. According to self-perception theory [9], when individuals perceive themselves as fully in control of their actions, they are more likely to attribute their decisions to internal reasoning rather than external influence. This increased self-reliance could, paradoxically, reduce their intention to incorporate the avatar's estimate. For example, one participant described this feeling as (translated to English): *"Even though I was no longer in the avatar when answering the question, I still believed that I was that person just now, so I believed in my own cognition more."* Future work could explore this by measuring social influence intention [11], or by introducing situational cues that modulate perceived authority or credibility of the avatar to better understand the interaction between agency and cognitive perspective-taking.

5 Discussion

This work used a controlled experiment design to investigate how spatial presence and agency, two fundamental qualities of VR, affect cognitive and affective perspective-taking in VR. Across two experiments, we manipulated spatial presence (via environmental detail: 2D vs. 3D scenarios) and agency (via temporal synchrony: high vs. low control, with or without a task), and measured cognitive and affective perspective-taking. Taken together, the results showed a clear dissociation between the two types of perspective-taking: higher spatial presence (Figure 7) and agency (Figure 12) significantly enhanced affective perspective-taking, but neither spatial presence (Figure 6) nor agency (Figure 11) had a significant effect on cognitive perspective-taking. Notably, the self-avatar overlap score was highest in the high-agency-without-task condition, while there was no significant difference between the high-agency-with-task and low-agency-without-task conditions. This suggests that cognitive load seems to play a key role. Therefore, we suggest that when designing tasks in VR scenarios, we should not pursue high fidelity of the scene as a default, but should consider the balance between scene and task complexity. This result is contrary to the previous intuition that a more realistic and detailed VR environment will inevitably lead to stronger perspective-taking effects.

5.1 Implications of Spatial Presence and Agency on VR Perspective-Taking

Our studies show that both spatial presence and agency enhance participants' affective perspective-taking, suggesting that they are both effective ways to enhance self-other overlap in VR.

For spatial presence, we primarily affected the self-location aspect of Wirth et al. [108]'s spatial presence model. We manipulated only the person's surroundings and therefore did not expect any effects on the possible actions part of the model, which is why we did not measure it. Participants reported a stronger feeling of "being there" in the mediated environment, as measured by the item " I felt like I was actually there in the environment of the presentation." [46]. However, because of the subjective nature of presence and the difficulties of manipulating it in a meaningful way [88], we added analyses on the median split scores of participants in addition to the ones comparing the values in the respective environments. Therefore, our results are based on the actual experience of spatial presence rather than the intended environmental manipulation. We chose spatial presence as defined by Wirth et al. [108] because of the key difference of VR compared to other, 2D, environments, as we were interested if the spatial component of VR would make a difference on perspective-taking outcomes [49]. However, because our study operationalizes cognitive and affective perspective-taking as distinct constructs grounded in psychological theory, our findings should be interpreted within the broader context of perspective-taking research.

In comparison, our agency manipulation targeted the control component of embodiment as described by Kilteni et al. [63], that is, the extent to which users experience themselves as the source of the avatar's movements rather than as passive observers.

However, it is important to note that while presence and agency both increased affective perspective-taking, neither was sufficient to shift cognitive perspective-taking as operationalized by the anchoring task. In other words, participants felt more similar to their avatar, but this subjective merging did not translate into greater susceptibility to the avatar's estimate in a decision task. This finding complicates common assumptions that deeper embodiment necessarily leads to greater adoption of an avatar's perspective in all cognitive or social tasks. Instead, it suggests that embodiment and perspective-taking may be partly dissociable processes in VR: *users can "feel similar" to the avatar without necessarily "thinking like" the avatar.*

While our work's primary goal was not to study embodiment as a dedicated goal, our findings draw connections to embodiment research and theory (e.g., because agency is a foundational pillar of embodiment [59, 90]). However, because perspective-taking and embodiment are related but separate constructs we discuss their relationship in the next paragraph.

5.2 The Relationship Between VR Perspective-Taking and Embodiment

The term "perspective-taking" has different interpretations in the VR and psychology literature. A we discuss in Section 2.1, psychology typically differentiates spatial, cognitive, and affective perspective-taking [25, 36], where the latter two are often summarized under psychological perspective-taking. Generally, perspective-taking is understood as the ability to understand "another person's thoughts, feelings, and inner mental states" [26]. In VR research, the term can also include aspects related to the way a person experiences a virtual environment visually. For example, some researchers use it to describe the way a person sees a VR experience. For instance, Hoppe et al. [51] present a perspective continuum that combines aspects of embodiment (e.g., within-character versus out-of-character) with aspects of view point (e.g., within-character versus out-of-character viewpoint). Aitamurto et al. [2] consider perspective-taking as taking the perspective of a person in a story. They investigate if a person sees a story shown in a video as an observer or as an actor. In their work, perspective describes the viewpoint of a person in a 180° or 360° video, and while the person can see a body, they are not embodied.

Related to this, embodiment is often discussed alongside VR perspective-taking but captures a different mechanism. Embodiment concerns the extent to which users experience a virtual body as their own, making it primarily a self-focused process that does not inherently require representing another person's mental states. This contrasts with psychological definitions of perspective-taking, which emphasize understanding another individual's thoughts or feelings [26]. One could ask if embodiment is an indicator of affective perspective-taking, because stronger embodiment is associated with identification with an avatar [73]. From a psychological perspective, however, this move risks conflating a self-focused bodily experience (embodiment) with an other-focused process of taking another person's perspective. Making this conceptual distinction explicit can help avoid conflating self-focused embodied experiences with other-centered psychological perspective-taking, thereby improving clarity and comparability across research in this area.

Yong et al. [112] specifically refer to embodied perspective-taking. In their study, participants watched a 360° video of a conversation between two partners from the other partner's point of view. So while participants saw the video from another view, they did not actually embody an avatar, which is often implicitly implied when using "embodied" perspective-taking. For example, other works refer to embodied perspective-taking as embodying a self-avatar of a fellow student [104] or embodying a self-avatar resembling an older person [76]. In that way VR perspective-taking can include embodiment of an avatar but does not have to. Simply watching a video in a VR environment is also considered VR perspective-taking, and described as "embodied" perspective-taking by some [112], others reserve the term for the embodiment of a self-avatar [76, 104]. However, they way a person "psychologically takes a perspective" is not included in those discussions (e.g., the way a person is instructed in a perspective-taking task, such as imagining oneself in a specific situation [7]). Our contributions are within the psychological understanding of perspective-taking, where our study reports how two VR qualities influence psychological perspective-taking outcomes. We investigated the influence of spatial presence and agency on psychological outcomes but were not interested in embodiment effects in our studies but found it useful to explain the difference between the two to advance research on VR perspective-taking.

The polysemous meaning of the term "perspective-taking" makes it challenging to synthesize research in the area of VR perspective-taking. However, we argue that this makes research aiming to bridge aspects of psychological perspective-taking with studies of how people experience VR (e.g., through presence and agency) particularly interesting and valuable. Such integration can help us understand not only how VR perspective-taking influences VR-specific qualities (e.g., embodiment) but also how it affects psychological constructs of perspective-taking. Our study represents an initial step toward examining the impact of VR-specific qualities on psychological perspective-taking outcomes in a controlled setting. We encourage future research to further explore the technical and perceptual factors that make VR "special" for perspective-taking, in particular in studying its effect on psychological perspective-taking outcomes.

5.3 Explaining the Dissociation Between Affective and Cognitive Perspective-Taking

First, our results suggest that although cognitive perspective-taking and affective perspective-taking are often mentioned together by researchers as social perspective-taking, they may be differently susceptible in their influences. This dissociation is consistent with prior neuroscientific findings suggesting that the two processes rely on partly distinct neural mechanisms, with cognitive perspective-taking more associated with the dorsomedial prefrontal cortex and affective perspective-taking more associated with the ventromedial prefrontal cortex [22].

Cognitive perspective-taking relies on higher-level, more effortful cognitive processes such as reasoning, memory retrieval, and executive function [28]. Our results suggest that these processes are not susceptible to environmental cues alone. The observed null effects may be because cognitive adjustments require explicit, controlled strategies that are relatively isolated from "immersive" or "embodied" experiences, or because participants maintained self-other boundaries during the reasoning task even when they felt more emotionally aligned. This dissociation is consistent with models that separate emotional "integration" from cognitive "simulation" or reasoning [21].

In addition, the elevated cognitive load imposed by immersive VR environments may have further limited participants' capacity for such controlled cognitive processes. Prior research has demonstrated that high-immersion VR, compared to low-immersion or desktop-based environments, substantially increases cognitive load, as users must continuously process complex multisensory input, manage spatial orientation, and interact with rich virtual stimuli [69, 77]. This heightened cognitive load may deplete the attentional and executive resources necessary for deliberate anchor adjustment or perspective-taking. As a result, even though immersive cues may enhance affective engagement, they can paradoxically reduce participants' ability to engage in the kind of effortful, controlled reasoning required for cognitive perspective-taking. This may help explain why, in our study, increased presence or agency did not translate into stronger cognitive alignment with the avatar.

Another possible explanation for the lack of significant cognitive perspective-taking effects relates to the "manipulation threshold" discussed in Experiment 1. Although both manipulations (presence in Experiment 1 and agency in Experiment 2) were validated and produced statistically significant changes in participants' subjective spatial presence or sense of agency, the absolute of these changes may have been insufficient to elicit measurable shifts in cognitive perspective-taking. For example, the difference in mean presence scores between the 3D and 2D conditions was relatively modest (3D: $M = 5.25$, $SD = 1.28$; 2D: $M = 4.84$, $SD = 1.45$), possibly failing to surpass a critical threshold needed to influence deeper cognitive processes. Similarly, as shown in recent work by Iarygina et al. [55], although the delay significantly affected participants' agency, it was not enough to break the sense of embodiment. Thus, it is possible that our presence and agency manipulations, while effective at enhancing affective engagement, were not strong enough to impact the more robust and deliberative mechanisms underlying cognitive perspective-taking.

5.4 Role of Task Demands and Cognitive Load

The experimental conditions of whether or not there was a task seemed to provide further explanation for the cognitive load argument. Specifically, while the higher autonomy condition resulted in higher affective perspective-taking overall, the *high no-task* condition obtained the highest scores, even exceeding the *high with-task* condition. This finding suggests that the lack of a goal-oriented task may enable participants to focus more fully on their embodied experience and sense of autonomy, thereby enhancing affective perspective-taking. Conversely, even in the high autonomy condition, adding a task seemed to slightly reduce subjective self-avatar overlap, perhaps due to increased cognitive demands. When cognitive resources are partially consumed by a task, there are fewer resources available for internal reflection and self-other integration. Therefore, a highly immersive, task-free environment may actually be most conducive to maximizing the fit between self and avatar because it allows participants to fully focus on their sense of embodiment without the distraction or cognitive resource fragmentation brought about by concurrent tasks.

5.5 Contributions and Implications

Taken together, our findings make several key contributions to the understanding of how spatial presence and agency influence VR perspective-taking. Firstly, by studying their main effects in two separate experiments and by incorporating both task and no-task conditions, our study provides evidence that affective and cognitive forms of perspective-taking are affected in different ways by immersive VR qualities. This extends previous literature by showing that while increased presence and agency reliably enhance subjective feelings of self-avatar overlap (affective perspective-taking), these changes do not necessarily translate into greater cognitive alignment or decision-making influence (cognitive perspective-taking).

Secondly, our results highlight the critical role of cognitive load in self–avatar relationships. We show that minimizing external task demands in highly immersive environments can further enhance self–avatar overlap, likely by freeing up cognitive resources for affective integration. Conversely, even modest increases in cognitive load (such as adding a task) can diminish this effect, pointing to a practical design consideration for VR applications aiming to maximize empathetic responses or self–other merging.

Applications in VR and Beyond. Our findings suggest important directions for the design of VR experiences across a variety of domains. In therapeutic or clinical settings, for example, VR systems that facilitate high presence and agency in a task-free environment may be most effective for interventions targeting empathy, self-concept, or emotion regulation, such as for individuals with social anxiety disorder. Similarly, in educational or training applications, balancing the degree of user autonomy, presence, and cognitive load could help tailor experiences for either affective learning (where embodiment and identification are crucial) or analytical reasoning (where excessive presence may in fact be detrimental).

For collaborative and social VR, our results indicate that optimizing conditions for affective resonance, such as minimizing unnecessary tasks and maximizing agency, may help foster empathy, trust, or group cohesion, especially in contexts where shared identity or perspective-taking are desired outcomes. Additionally, where independent reasoning and reduced susceptibility to social influence are important (e.g., decision-making or negotiation scenarios), designers may wish to intentionally increase cognitive load or task focus to maintain boundaries between self and avatar.

Overall, by clarifying the distinct mechanisms underlying affective and cognitive perspective-taking in VR, and by demonstrating how these processes can be independently modulated through design choices, our work offers actionable insights for researchers and practitioners seeking to leverage the unique affordances of immersive technologies for psychological, educational, or social outcomes.

5.6 Limitations and Future Work

While our work has systematically examined the independent effects of spatial presence and agency on cognitive and affective perspective-taking in VR, we acknowledge several limitations.

5.6.1 Limitations. First, we acknowledge potentially limited manipulation strength. Although both presence and agency manipulations achieved significant subjective effects, the absolute differences, especially for spatial presence, were relatively modest (*3D:* $M = 5.25$, *2D:* $M = 4.84$). Identifying an effective, meaningful, and justified way to manipulate presence remains a broader challenge in VR research. As indicated in several synthesis articles (e.g., by Xiao et al. [111] or Skarbez et al. [88]), presence is a multi-dimensional construct and it remains difficult to ensure detailed measures of its sub components. We chose spatial presence as a presence construct that is related to the space, as this is often mentioned in psychology perspective-taking studies as a major advantage of VR. We aimed to induce different levels of spatial presence to be able to measure a relationship to perspective-taking outcomes. Because presence ratings are highly subjective, in our analysis, we also included a median-split analysis to focus on the perceived presence rather than the environmental manipulation. It is possible that the manipulations did not reach the threshold required to elicit deeper changes in cognitive processes. Stronger or more ecologically varied manipulations (e.g., highly realistic social scenarios, more extreme agency disruptions) might produce different outcomes. Secondly, our primary measure of cognitive perspective-taking relied on a numerical anchoring paradigm, which, while widely used does not capture all facets of cognitive alignment in VR. Future work should combine anchoring with complementary, non-numerical measures of cognitive perspective-taking, such as counterfactual thinking. Additionally, in Experiment 1, we observed that some participants accepted and copied the avatar's anchor (i.e., they submitted the same value). This may have occurred because of low familiarity with the question or declining engagement during the task. These incidences potentially added noise to the anchoring results. Despite our improvements in Experiment 2 (i.e., participants could not enter the same value as the avatar), anchoring effects were still not significant. This pattern suggests that our experimental manipulations had little impact on cognitive perspective-taking. Thirdly, to ensure the validity of our results, both experiments adopted a controlled laboratory design with multiple repetitions per condition (10 trials per condition in Experiment 1 and 6 in Experiment 2). To mitigate the risk that such repetition might reveal the study purpose to participants, we introduced a cover story framed as a "sign language learning" task. This framing justified the repeated trials (as repetition is natural when learning a new skill) and encouraged participants to move their fingers, which aligned with our manipulation of avatar agency. Nevertheless, this approach may not have fully engaged participants, and repeated exposure may still have introduced demand characteristics despite our efforts. Fourth, although our paradigm is intentionally designed to be scenario-independent to avoid confounds tied to specific narratives, this design choice inevitably reduces ecological validity. Finally, our sample size was modest (23 and 24 participants, respectively), which may limit statistical power for detecting subtle cognitive effects. While our power analysis suggested sufficient sensitivity for medium effects, future work with larger and more diverse samples could improve robustness and generalizability.

5.6.2 Future Work. Based on these limitations, we suggest several directions for future research. Future studies should test stronger presence (e.g., photorealistic avatars) and agency manipulations (e.g., full loss of control) to better detect cognitive shifts. To address the limitations of the anchoring paradigm, future research could incorporate additional behavioral, physiological, or process-tracing measures of cognitive perspective-taking, such as theory-of-mind tasks, gaze alignment, or neural correlates, alongside self-report self-avatar overlap. This multi-method approach may capture subtler effects missed by the current design. Investigating how repeated or prolonged VR perspective-taking interventions influence both affective and cognitive alignment over time is essential. Longitudinal designs could test whether initial increases in perceived overlap or empathy translate into lasting changes in attitudes or social behavior. Expanding the participant pool to include individuals from different age groups, cultures, or levels of VR familiarity can test the robustness and external validity of the observed effects. This could also address individual difference moderators, such as baseline empathy. Finally, embedding perspective-taking tasks in more realistic, emotionally, or morally charged scenarios, such as social dilemmas, negotiations, or simulated discrimination experiences, may increase sensitivity to both affective and cognitive changes and better simulate real-world applications.

By addressing these points, future work can build on our studies' controlled paradigm to identify the boundary conditions and mechanisms through which VR qualities shape different facets of perspective-taking, and inform the design of more impactful and generalizable VR-based interventions.

6 Conclusion

This paper investigated how two key qualities of VR (spatial presence and agency) influence cognitive and affective perspective-taking. Across two experiments, we found that manipulating spatial presence through environmental realism (3D vs. 2D) and agency through control responsiveness (synchronous vs. delayed avatar motion) significantly affected participants' subjective experience of spatial presence and agency. High presence environments increased participants' subjective experience of spatial presence, while high agency conditions led to greater perceived control over the avatar. Importantly, these qualities had a consistent effect on participants' perceived self-avatar overlap: both higher presence (Experiment 1) and higher agency (Experiment 2) increased the degree to which participants felt similar to the avatar. However, contrary to our hypotheses, neither spatial presence nor agency significantly enhanced cognitive perspective-taking as measured through an anchoring task: participants' numerical estimates did not reliably shift toward those of the avatar under high presence or high agency conditions. This discrepancy highlights a possible dissociation between affective and cognitive alignment in VR perspective-taking. These findings suggest that while embodiment fosters avatar–self merging, it does not automatically translate to deeper cognitive perspective-taking. Our work systematically examines how spatial presence and agency influence different levels of perspective-taking, showing that both qualities enhance affective perspective-taking but do not produce corresponding improvements in cognitive perspective-taking.

Acknowledgments

We would like to thank all user study participants and the reviewers for their insightful and valuable feedback. This research was supported by the Pioneer Centre for Artificial Intelligence, DNRF grant number P1.

References

[1] Sun Joo (Grace) Ahn, Joshua Bostick, Elise Ogle, Kristine L. Nowak, Kara T. McGillicuddy, and Jeremy N. Bailenson. 2016. Experiencing Nature: Embodying Animals in Immersive Virtual Environments Increases Inclusion of Nature in Self and Involvement with Nature. *Journal of Computer-Mediated Communication* 21, 6 (2016), 399–419. https://doi.org/10.1111/jcc4.12173

[2] Tanja Aitamurto, Shuo Zhou, Sukolsak Sakshuwong, Jorge Saldivar, Yasamin Sadeghi, and Amy Tran. 2018. Sense of Presence, Attitude Change, Perspective-Taking and Usability in First-Person Split-Sphere 360° Video. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3174119

[3] Arthur Aron, Elaine N Aron, and Danny Smollan. 1992. Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology* 63, 4 (1992), 596. https://doi.org/10.1037/0022-3514.63.4.596

[4] Domna Banakou, Richard Groten, and Mel Slater. 2013. Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences* 110, 31 (July 2013), 12846–12851. https://doi.org/10.1073/pnas.1306779110

[5] Domna Banakou, Parasuram D. Hanumanthu, and Mel Slater. 2016. Virtual Embodiment of White People in a Black Virtual Body Leads to a Sustained Reduction in Their Implicit Racial Bias. *Frontiers in Human Neuroscience* 10 (2016), 601. https://doi.org/10.3389/fnhum.2016.00601

[6] Simon Baron-Cohen, Alan M Leslie, and Uta Frith. 1985. Does the autistic child have a "theory of mind"? *Cognition* 21, 1 (1985), 37–46. https://doi.org/10.1016/0010-0277(85)90022-8

[7] C Daniel Batson, Shannon Early, and Giovanni Salvarani. 1997. Perspective Taking: Imagining How Another Feels Versus Imaging How You Would Feel. *Personality and social psychology bulletin* 23, 7 (1997), 751–758. https://doi.org/10.1177/0146167297237008

[8] Dan Batson. 2018. Empathy, Altruism, and Helping: Conceptual Distinctions, Empirical Relations. In *Forms of Fellow Feeling: Empathy, Sympathy, Concern and Moral Agency*, Neil Roughley and Thomas Schramme (Eds.). Cambridge University Press, Cambridge, 59–77.

[9] Daryl J Bem. 1972. Self-perception theory. In *Advances in experimental social psychology*. Vol. 6. Elsevier, 1–62. https://doi.org/10.1016/S0065-2601(08)60024-6

[10] Jim Blascovich, Jack Loomis, Andrew C Beall, Kimberly R Swinth, Crystal L Hoyt, and Jeremy N Bailenson. 2002. TARGET ARTICLE: Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology. *Psychological inquiry* 13, 2 (2002), 103–124. https://doi.org/10.1207/S15327965PLI1302_01

[11] Gerd Bohner, Sabine Einwiller, Hans-Peter Erb, and Frank Siebler. 2003. When small means comfortable: Relations between product attributes in two-sided advertising. *Journal of Consumer Psychology* 13, 4 (2003), 454–463. https://doi.org/10.1207/S15327663JCP1304_12

[12] Zahra Borhani and Francisco R. Ortega. 2023. A Virtual Reality System for Gender Swapping to Increase Empathy Against Stereotype Threats in Computer Science Job Interviews. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 648–649. https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00138

[13] Gretchen B. Chapman and Eric J. Johnson. 2002. Incorporating the irrelevant: Anchors in judgments of belief and value. In *Heuristics and Biases: The Psychology of Intuitive Judgment*, Thomas Gilovich, Dale Griffin, and Daniel Kahneman (Eds.). Cambridge University Press, Cambridge, 120–138. https://doi.org/10.1017/CBO9780511808098.008

[14] Tanvir Irfan Chowdhury, Sharif Mohammad Shahnewaz Ferdous, and John Quarles. 2019. VR disability simulation reduces implicit bias towards persons with disabilities. *IEEE transactions on visualization and computer graphics* 27, 6 (2019), 3079–3090. https://doi.org/10.1109/TVCG.2019.2958332

[15] María Jimena Clavel Vázquez and Adriana Clavel-Vázquez. 2023. Robustly embodied imagination and the limits of perspective-taking. *Philosophical Studies* 180, 4 (2023), 1395–1420. https://doi.org/10.1007/s11098-023-01936-4

[16] Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2 ed.). Routledge, New York. 567 pages. https://doi.org/10.4324/9780203771587

[17] Olivier Corneille and Peter Lush. 2023. Sixty Years After Orne's American Psychologist Article: A Conceptual Framework for Subjective Experiences Elicited by Demand Characteristics. *Personality and Social Psychology Review* 27, 1 (2023), 83–101. https://doi.org/10.1177/10888683221104368

[18] Mark H. Davis. 1983. Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach. *Journal of Personality and Social Psychology*

44, 1 (1983), 113–126. https://doi.org/10.1037/0022-3514.44.1.113

[19] Mark H Davis, Laura Conklin, Amy Smith, and Carol Luce. 1996. Effect of perspective taking on the cognitive representation of persons: a merging of self and other. *Journal of personality and social psychology* 70, 4 (1996), 713. https://doi.org/10.1037/0022-3514.70.4.713

[20] Jean Decety and Philip L Jackson. 2004. The functional architecture of human empathy. *Behavioral and cognitive neuroscience reviews* 3, 2 (2004), 71–100. https://doi.org/10.1177/1534582304267187

[21] Jean Decety and Philip L Jackson. 2006. A social-neuroscience perspective on empathy. *Current directions in psychological science* 15, 2 (2006), 54–58. https://doi.org/10.1111/j.0963-7214.2006.00406.x

[22] Bryan T. Denny, Hedy Kober, Tor D. Wager, and Kevin N. Ochsner. 2012. A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *Journal of Cognitive Neuroscience* 24, 8 (2012), 1742–1752. https://doi.org/10.1162/jocn_a_00233

[23] Nyaz Didehbani, Tandra Allen, Michelle Kandalaft, Daniel Krawczyk, and Sandra Chapman. 2016. Virtual Reality Social Cognition Training for children with high functioning autism. *Comput. Hum. Behav.* 62, C (Sept. 2016), 703–711. https://doi.org/10.1016/j.chb.2016.04.033

[24] Nancy Eisenberg and Paul A Miller. 1987. The relation of empathy to prosocial and related behaviors. *Psychological bulletin* 101, 1 (1987), 91–119. https://doi.org/10.1037/0033-2909.101.1.91

[25] R. D. Enright and D. K. Lapsley. 1980. Social Role-taking: A Review of the Constructs, Measures, and Measurement Properties. *Review of Educational Research* 50, 4 (1980), 647–674. https://doi.org/10.3102/00346543050004647

[26] Nicholas Epley and Thomas Gilovich. 2001. Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological science* 12, 5 (2001), 391–396. https://doi.org/10.1111/1467-9280.00372

[27] Nicholas Epley and Thomas Gilovich. 2005. When effortful thinking influences judgmental anchoring: differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of behavioral decision making* 18, 3 (2005), 199–212. https://doi.org/10.1002/bdm.495

[28] Nicholas Epley and Thomas Gilovich. 2006. The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science* 17, 4 (2006), 311–318. https://doi.org/10.1111/j.1467-9280.2006.01704.x

[29] Nicholas Epley, Boaz Keysar, Leaf Van Boven, and Thomas Gilovich. 2004. Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology* 87, 3 (2004), 327–339. https://doi.org/10.1037/0022-3514.87.3.327

[30] Thorsten M. Erle and Friederike Funk. 2022. Visuospatial and affective perspective-taking: Similarities and differences. *Social Psychology* 53, 5 (2022), 315–326. https://doi.org/10.1027/1864-9335/a000504

[31] Thorsten M Erle and Sascha Topolinski. 2017. The grounded nature of psychological perspective-taking. *Journal of personality and social psychology* 112, 5 (2017), 683–695. https://doi.org/10.1037/pspa0000081

[32] Ronald Aylmer Fisher. 1971. *The design of experiments*. Springer.

[33] John H Flavell. 1977. The development of knowledge about visual perception.. In *Nebraska Symposium on Motivation. Nebraska Symposium on Motivation*, Vol. 25. 43–76.

[34] John H Flavell. 2004. Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly (1982-)* (2004), 274–290. https://doi.org/10.1353/mpq.2004.0018

[35] John H Flavell, Barbara A Everett, Karen Croft, and Eleanor R Flavell. 1981. Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental psychology* 17, 1 (1981), 99–103. https://doi.org/10.1037/0012-1649.17.1.99

[36] M. E. Ford. 1979. The construct validity of egocentrism. *Psychological Bulletin* 86, 6 (1979), 1169–1188. https://doi.org/10.1037/0033-2909.86.6.1169

[37] Wallace V Friesen and Paul Ekman. 1976. *Pictures of facial affect*. Consulting psychologists press.

[38] Chris D Frith, Sarah-Jayne Blakemore, and Daniel M Wolpert. 2000. Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain research reviews* 31, 2-3 (2000), 357–363. https://doi.org/10.1016/s0165-0173(99)00052-1

[39] Adam D. Galinsky, Gillian Ku, and Cynthia S. Wang. 2005. Perspective-Taking and Self-Other Overlap: Fostering Social Bonds and Facilitating Social Coordination. *Group Processes & Intergroup Relations* 8, 2 (2005), 109–124. https://doi.org/10.1177/1368430205051060

[40] Adam D. Galinsky and Thomas Mussweiler. 2001. First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology* 81, 4 (2001), 657–669. https://doi.org/10.1037/0022-3514.81.4.657

[41] Mar Gonzalez-Franco, Anthony Steed, Christopher C Berger, and Ana Tajadura-Jiménez. 2024. The impact of first-person avatar customization on embodiment in immersive virtual reality. *Frontiers in Virtual Reality* 5 (2024), 1436752. https://doi.org/10.3389/frvir.2024.1436752

[42] Alison Gopnik and Virginia Slaughter. 1991. Young Children's Understanding of Changes in Their Mental States. *Child development* 62, 1 (1991), 98–110. https://doi.org/10.2307/1130707

[43] Anthony G Greenwald. 1976. Within-subjects designs: To use or not to use? *Psychological Bulletin* 83, 2 (1976), 314–320. https://doi.org/10.1037/0033-2909.83.2.314

[44] Victoria Groom, Jeremy N Bailenson, and Clifford Nass. 2009. The influence of racial embodiment on racial bias in immersive virtual environments. *Social Influence* 4, 3 (2009), 231–248. https://doi.org/10.1080/15534510802643750

[45] Catherine Hamilton-Giachritsis, Domna Banakou, Manuela Garcia Quiroga, Christos Giachritsis, and Mel Slater. 2018. Reducing risk and improving maternal perspective-taking and empathy using virtual embodiment. *Scientific Reports* 8, 1 (Feb. 2018), 2975. https://doi.org/10.1038/s41598-018-21036-2

[46] Tilo Hartmann, Werner Wirth, Holger Schramm, Christoph Klimmt, Peter Vorderer, André Gysbers, Saskia Böcking, Niklas Ravaja, Jari Laarni, Timo Saari, et al. 2015. The spatial presence experience scale (SPES). *Journal of Media Psychology* (2015).

[47] Béatrice S Hasler, Bernhard Spanlang, and Mel Slater. 2017. Virtual race transformation reverses racial in-group bias. *PloS one* 12, 4 (2017), e0174965. https://doi.org/10.1371/journal.pone.0174965

[48] Meghan L Healey and Murray Grossman. 2018. Cognitive and Affective Perspective-Taking: Evidence for Shared and Dissociable Anatomical Substrates. *Frontiers in neurology* 9 (2018), 491. https://doi.org/10.3389/fneur.2018.00491

[49] Fernanda Herrera, Jeremy N. Bailenson, Erika Weisz, Elise Ogle, and Jamil Zaki. 2018. Building Long-Term Empathy: A Large-Scale Comparison of Traditional and Virtual Reality Perspective-Taking. *PLOS ONE* 13, 10 (2018), e0204494. https://doi.org/10.1371/journal.pone.0204494

[50] Hal E. Hershfield, Daniel G. Goldstein, William F. Sharpe, Jesse Fox, Leo Yeykelis, Laura L. Carstensen, and Jeremy N. Bailenson. 2011. Increasing Saving Behavior Through Age-Progressed Renderings of the Future Self. *Journal of Marketing Research* 48, SPL (2011), S23–S37. https://doi.org/10.1509/jmkr.48.SPL.S23

[51] Matthias Hoppe, Andrea Baumann, Patrick Chofor Tamunjoh, Tonja-Katrin Machulla, Paweł W. Woźniak, Albrecht Schmidt, and Robin Welsch. 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* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 360, 13 pages. https://doi.org/10.1145/3491102.3517447

[52] Patricia Howlin, Simon Baron-Cohen, and Julie A Hadwin. 1999. *Teaching children with autism to mind-read: A practical guide for teachers and parents*. John Wiley & Sons.

[53] Judy Dunn Claire Hughes. 1998. Young Children's Understanding of Emotions within Close Relationships. *Cognition and Emotion* 12, 2 (1998), 171–190. https://doi.org/10.1080/026999398379709

[54] Olga Iarygina, Kasper Hornbæk, and Aske Mottelson. 2025. Demand characteristics in human–computer experiments. *Int. J. Hum.-Comput. Stud.* 193, C (Jan. 2025), 14 pages. https://doi.org/10.1016/j.ijhcs.2024.103379

[55] Olga Iarygina, Kasper Hornbæk, and Aske Mottelson. 2025. Does Random Movements mean Random Results? Why Asynchrony in Experiments on Body Ownership does not Work as Intended. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 369, 19 pages. https://doi.org/10.1145/3706598.3713506

[56] Carley Jans, Fiona Bogossian, Patrea Andersen, and Tracy Levett-Jones. 2023. Examining the impact of virtual reality on clinical decision making–An integrative review. *Nurse Education Today* 125 (2023), 105767. https://doi.org/10.1016/j.nedt.2023.105767

[57] Marc Jeannerod. 2001. Neural Simulation of Action: A Unifying Mechanism for Motor Cognition. *NeuroImage* 14, 1 Pt 2 (Jul 2001), S103–S109. https://doi.org/10.1006/nimg.2001.0832

[58] Crescent Jicol, Hoi Ying Cheng, Karin Petrini, and Eamonn O'Neill. 2023. A predictive model for understanding the role of emotion for the formation of presence in virtual reality. *Plos one* 18, 3 (2023), e0280390. https://doi.org/10.1371/journal.pone.0280390

[59] Andreas Kalckert and H Henrik Ehrsson. 2012. Moving a rubber hand that feels like your own: a dissociation of ownership and agency. *Frontiers in Human Neuroscience* 6 (2012), 40. https://doi.org/10.3389/fnhum.2012.00040

[60] Michelle R Kandalaft, Nyaz Didehbani, Daniel C Krawczyk, Tandra T Allen, and Sandra B Chapman. 2013. Virtual reality social cognition training for young adults with high-functioning autism. *Journal of autism and developmental disorders* 43, 1 (2013), 34–44. https://doi.org/10.1007/s10803-012-1544-6

[61] Klaus Kessler and Hannah Rutherford. 2010. The two forms of visuo-spatial perspective taking are differently embodied and subserve different spatial prepositions. *Frontiers in psychology* 1 (2010), 213. https://doi.org/10.3389/fpsyg.2010.00213

[62] Boaz Keysar, Dale J Barr, Jennifer A Balin, and Jason S Brauner. 2000. Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological science* 11, 1 (2000), 32–38. https://doi.org/stable/40063492

[63] Konstantina Kilteni, Raphaela Groten, and Mel Slater. 2012. The Sense of Embodiment in Virtual Reality. *Presence: Teleoper. Virtual Environ.* 21, 4 (Dec. 2012), 373–387. https://doi.org/10.1162/PRES_a_00124

[64] Claus Lamm, C Daniel Batson, and Jean Decety. 2007. The Neural Substrate of Human Empathy: Effects of Perspective-taking and Cognitive Appraisal. *Journal of cognitive neuroscience* 19, 1 (2007), 42–58. https://doi.org/10.1162/jocn.2007.19.1.42

[65] Kwan Min Lee. 2004. Presence, explicated. *Communication Theory* 14, 1 (2004), 27–50. https://doi.org/10.1111/j.1468-2885.2004.tb00302.x

[66] Peter Lush. 2020. Demand characteristics confound the rubber hand illusion. *Collabra: Psychology* 6, 1 (2020), 22. https://doi.org/10.1525/collabra.325

[67] Lara Maister, Natalie Sebanz, Günther Knoblich, and Manos Tsakiris. 2013. Experiencing ownership over a dark-skinned body reduces implicit racial bias. *Cognition* 128, 2 (2013), 170–178. https://doi.org/10.1016/j.cognition.2013.04.002

[68] Lara Maister, Mel Slater, Maria V Sanchez-Vives, and Manos Tsakiris. 2015. Changing bodies changes minds: owning another body affects social cognition. *Trends in cognitive sciences* 19, 1 (2015), 6–12. https://doi.org/10.1016/j.tics.2014.11.001

[69] Guido Makransky, Thomas S Terkildsen, and Richard E Mayer. 2019. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and instruction* 60 (2019), 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

[70] Marco Marinucci and Paolo Riva. 2025. Homelessness in Virtual Reality: Experiencing Social Exclusion Improves Attitudes and Behaviors Toward Unhoused People. *Cyberpsychology, Behavior, and Social Networking* 28, 1 (2025), 65–71. https://doi.org/10.1089/cyber.2024.0027 arXiv:https://doi.org/10.1089/cyber.2024.0027 PMID: 39512124.

[71] George Herbert Mead. 1934. Mind, self, and society from the standpoint of a social behaviorist. (1934).

[72] Henrike Moll and Michael Tomasello. 2006. Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology* 24, 3 (2006), 603–613. https://doi.org/10.1348/026151005X55370DigitalObjectIdentifier(DOI)

[73] Aske Mottelson, Andreea Muresan, Kasper Hornbæk, and Guido Makransky. 2023. A Systematic Review and Meta-analysis of the Effectiveness of Body Ownership Illusions in Virtual Reality. *ACM Trans. Comput.-Hum. Interact.* 30, 5, Article 76 (Sept. 2023), 42 pages. https://doi.org/10.1145/3590767

[74] Jochen Müsseler, Sophia von Salm-Hoogstraeten, and Christian Böffel. 2022. Perspective taking and avatar-self merging. *Frontiers in Psychology* 13 (2022), 714464. https://doi.org/10.3389/fpsyg.2022.714464

[75] Thomas Mussweiler and Fritz Strack. 2001. The Semantics of Anchoring. *Organizational Behavior and Human Decision Processes* 86, 2 (2001), 234–255. https://doi.org/10.1006/obhd.2001.2954

[76] Soo Youn Oh, Jeremy Bailenson, Erika Weisz, and Jamil Zaki. 2016. Virtually old: Embodied perspective taking and the reduction of ageism under threat. *Computers in Human Behavior* 60 (2016), 398–410. https://doi.org/10.1016/j.chb.2016.02.007

[77] Jocelyn Parong and Richard E Mayer. 2018. Learning science in immersive virtual reality. *Journal of educational psychology* 110, 6 (2018), 785–797. https://doi.org/10.1037/edu0000241

[78] Tabitha C. Peck, Sofia Seinfeld, Salvatore M. Aglioti, and Mel Slater. 2013. Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition* 22, 3 (2013), 779–787. https://doi.org/10.1016/j.concog.2013.04.016

[79] Jean Piaget, Baerbel Inhelder, F. J. Langdon, and J. L. Lunzer. 1957. The Child's Conception of Space. *British Journal of Educational Studies* 5, 2 (1957), 187–189. https://doi.org/10.2307/3118882

[80] David Premack and Guy Woodruff. 1978. Does the chimpanzee have a theory of mind? *Behavioral and brain sciences* 1, 4 (1978), 515–526. https://doi.org/10.1017/S0140525X00076512

[81] Timothy B. Rogers, Nicholas A. Kuiper, and William S. Kirker. 1977. Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology* 35, 9 (Sep 1977), 677–688. https://doi.org/10.1037//0022-3514.35.9.677

[82] Robert Rosenthal. 1976. Experimenter effects in behavioral research. (1976).

[83] Daniel Roth and Marc Erich Latoschik. 2020. Construction of the Virtual Embodiment Questionnaire (VEQ). *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (2020), 3546–3556. https://doi.org/10.1109/TVCG.2020.3023603

[84] Richard M. Ryan and Edward L. Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist* 55, 1 (2000), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

[85] Dana Samson, Ian A Apperly, Jason J Braithwaite, Benjamin J Andrews, and Sarah E Bodley Scott. 2010. Seeing it their way: evidence for rapid and involuntary computation of what other people see. *Journal of experimental psychology: human perception and performance* 36, 5 (2010), 1255. https://doi.org/10.1037/a0018729

[86] Catherine L Sebastian, Nathalie MG Fontaine, Geoffrey Bird, Sarah-Jayne Blakemore, Stephane A De Brito, Eamon JP McCrory, and Essi Viding. 2012. Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *Social cognitive and affective neuroscience* 7, 1 (2012), 53–63. https://doi.org/10.1093/scan/nsr023

[87] Roger N Shepard and Jacqueline Metzler. 1971. Mental rotation of three-dimensional objects. *Science* 171, 3972 (1971), 701–703. https://doi.org/10.1126/science.171.3972.701

[88] Richard Skarbez, Frederick P. Brooks, Jr., and Mary C. Whitton. 2017. A Survey of Presence and Related Concepts. *ACM Comput. Surv.* 50, 6, Article 96 (Nov. 2017), 39 pages. https://doi.org/10.1145/3134301

[89] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

[90] Mel Slater, Bernhard Spanlang, Maria V Sanchez-Vives, and Olaf Blanke. 2010. First person experience of body transfer in virtual reality. *PloS one* 5, 5 (2010), e10564. https://doi.org/10.1371/journal.pone.0010564

[91] Mel Slater and Sylvia Wilbur. 1997. A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 6, 6 (1997), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

[92] Jonathan Steuer et al. 1992. Defining virtual reality: Dimensions determining telepresence. *Journal of communication* 42, 4 (1992), 73–93. https://doi.org/10.1111/j.1460-2466.1992.tb00812.x

[93] Ezra Stotland. 1969. Exploratory Investigations of Empathy. *Advances in Experimental Social Psychology* 4 (1969), 271–314. https://doi.org/10.1016/S0065-2601(08)60080-5

[94] Cynthia S Symons and Blair T Johnson. 1997. The self-reference effect in memory: a meta-analysis. *Psychological bulletin* 121, 3 (1997), 371. https://doi.org/10.1037/0033-2909.121.3.371

[95] Ana Tajadura-Jiménez, Domna Banakou, Nadia Bianchi-Berthouze, and Mel Slater. 2017. Embodiment in a Child-Like Talking Virtual Body Influences Object Size Perception, Self-Identification, and Subsequent Real Speaking. *Scientific Reports* 7, 1 (Aug. 2017), 9637. https://doi.org/10.1038/s41598-017-09497-3

[96] Henri Tajfel and John C Turner. 2004. The social identity theory of intergroup behavior. In *Political psychology*. Psychology Press, 276–293. https://doi.org/10.4324/9780203505984-16

[97] Diana I Tamir and Jason P Mitchell. 2010. Neural correlates of anchoring-and-adjustment during mentalizing. *Proceedings of the National Academy of Sciences* 107, 24 (2010), 10827–10832. https://doi.org/10.1073/pnas.1003242107

[98] Peter Thompson and David Burr. 2009. Visual aftereffects. *Current Biology* 19, 1 (2009), R11–R14.

[99] Michael Tomasello. 2009. *Why we cooperate.* MIT press. https://doi.org/10.1017/S0266267111000071

[100] Manos Tsakiris, Matthew R. Longo, and Patrick Haggard. 2010. Having a body versus moving your body: neural signatures of agency and body-ownership. *Neuropsychologia* 48, 9 (July 2010), 2740–2749. https://doi.org/10.1016/j.neuropsychologia.2010.05.021 Copyright (c) 2010 Elsevier Ltd. All rights reserved..

[101] Amos Tversky and Daniel Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science* 185, 4157 (1974), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

[102] Barbara Tversky and Bridgette Martin Hard. 2009. Embodied and disembodied cognition: Spatial perspective-taking. *Cognition* 110, 1 (2009), 124–129. https://doi.org/10.1016/j.cognition.2008.10.008

[103] Jean-Louis Van Gelder, Liza JM Cornet, Natascha P Zwalua, Esther CA Mertens, and Job van der Schalk. 2022. Interaction with the future self in virtual reality reduces self-defeating behavior in a sample of convicted offenders. *Scientific reports* 12, 1 (2022), 2254. https://doi.org/10.1038/s41598-022-06305-5

[104] Austin van Loon, Jeremy Bailenson, Jamil Zaki, Joshua Bostick, and Robb Willer. 2018. Virtual reality perspective-taking increases cognitive empathy for specific others. *PLOS ONE* 13, 8 (08 2018), 1–19. https://doi.org/10.1371/journal.pone.0202442

[105] Theresa K Vescio, Gretchen B Sechrist, and Matthew P Paolucci. 2003. Perspective taking and prejudice reduction: The mediational role of empathy arousal and situational attributions. *European journal of social psychology* 33, 4 (2003), 455–472. https://doi.org/10.1002/ejsp.163

[106] Henry M Wellman, David Cross, and Julanne Watson. 2001. Meta-analysis of theory-of-mind development: The truth about false belief. *Child development* 72, 3 (2001), 655–684. https://doi.org/10.1111/1467-8624.00304

[107] Henry M Wellman, Ann T Phillips, and Thomas Rodriguez. 2000. Young children's understanding of perception, desire, and emotion. *Child development* 71, 4 (2000), 895–912. https://doi.org/10.1111/1467-8624.00198

[108] Werner Wirth, Tilo Hartmann, Saskia Böcking, Peter Vorderer, Christoph Klimmt, Holger Schramm, Timo Saari, Jari Laarni, Niklas Ravaja, Feliz Ribeiro Gouveia, et al. 2007. A process model of the formation of spatial presence experiences. *Media psychology* 9, 3 (2007), 493–525. https://doi.org/10.1080/15213260701283079

[109] Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 3 (1998), 225–240. https:

//doi.org/10.1162/105474698565686
[110] Kyle M Woosnam et al. 2010. The inclusion of other in the self (IOS) scale. *Annals of Tourism Research* 37, 3 (2010), 857–860.
[111] Cleo Xiao, Difeng Yu, Kasper Hornbæk, and Joanna Bergström. 2025. A Concept at Work: A Review of Motivations, Operationalizations, and Conclusions in VR Research about Presence. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 1145, 21 pages. https://doi.org/10.1145/3706598.3714279
[112] Seraphina Yong, Leo Cui, Evan Suma Rosenberg, and Svetlana Yarosh. 2024. A Change of Scenery: Transformative Insights from Retrospective VR Embodied Perspective-Taking of Conflict With a Close Other. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 919, 18 pages. https://doi.org/10.1145/3613904.3642146
[113] Christopher You, Roshan Venkatakrishnan, Rohith Venkatakrishnan, Zhuoming Han, Benjamin Lok, and Tabitha Peck. 2024. A Sense of Urgency on the Sense of Agency: Challenges in Evaluating Agency and Embodiment in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 30, 11 (2024), 7172–7182. https://doi.org/10.1109/TVCG.2024.3456139
[114] Jamil Zaki and Kevin N Ochsner. 2012. The neuroscience of empathy: progress, pitfalls and promise. *Nature neuroscience* 15, 5 (2012), 675–680. https://doi.org/10.1038/nn.3085

A Full List of On-Screen Instructions and Questions Used in the Experiment

This appendix lists all instructions and questions presented to participants during the study. Questions were delivered inside the VR environment using a graphical user interface.

A.1 Experiment 1

A.1.1 General Instructions (shown in VR at the start of each trial).

> 'In this experience you will help this person learn letters from the American Sign Language alphabet. For this purpose, you will be embodied in the avatar you are seeing. This means that your movements will be aligned with the avatar's movement. If you are ready to start the experience, click 'start'.'

A.1.2 Presence Questions . Participants answered the following statements on a 7-point Likert scale (1 = fully disagree, 7 = fully agree):

- 'I feel that I am actually in this virtual environment'

A.1.3 Anchor Estimation Task. Participants were presented with a scenario-specific numerical estimate made by another person (the "anchor") and asked:

> *'Please read the following anchor question carefully:'*
> *e.g., 'The other person says [anchor]. What do you think?'*

They then verbally reported their numerical estimates to the experimenter.

A.1.4 Self-Avatar Overlap Rating.

> 'Do you think you are similar to this person (in any aspect, world view, values, cognition, etc.)?'

(7-point Likert scale: 1 = fully disagree, 7 = fully agree)

A.1.5 Continue Prompt.

> 'I am ready to the next trial. Please press Continue to proceed.'

A.2 Experiment 2

Experiment 2 included the same procedure and questions as Experiment 1, with two additions: a block of agency questions and a confidence rating.

A.2.1 Agency Questions (After Embodiment). Participants rated their experience of control and embodiment using the following four statements on a 7-point Likert scale (1 = fully, 7 = fully agree):

- 'The movements of the virtual body felt like they were my movements.'
- 'I felt like I was controlling the movements of the virtual body.'
- 'I felt like I was causing the movements of the virtual body.'
- 'The movements of the virtual body were in sync with my own movements.'

A.2.2 Confidence Rating (After Estimation).

> 'How confident are you in your answer?'

(7-point Likert scale: 1 = not confident at all, 7 = very confident)

A.2.3 Presence, Anchor, Confidence, and Self-Avatar Overlap Questions. The same set of presence questions, anchor estimation, confidence rating, and self-avatar overlap rating from Experiment 1 were reused in Experiment 2. See previous section for full wording.

A.3 Results

A.3.1 Exploring influences of Presence and Anchor Direction on Cognitive perspective-taking. In order to further explore how perceived presence ("PresenceGroup") and anchor direction ("AnchorGroup") might influence the absolute deviation from the anchor value, we employed both multiple linear regression and two-way ANOVA. The **multiple linear regression** allowed us to include interaction terms and continuous predictors, providing a flexible model to examine potential moderating effects. The **two-way ANOVA**, on the other hand, treated both factors as categorical and provided an interpretable framework for testing main and interaction effects with fewer assumptions.

We included both analyses to ensure our findings were not dependent on a single analytical approach. While the regression model captured nuance in the data, the ANOVA allowed for a straightforward examination of whether categorical groupings (e.g., high/low presence and high/low anchor) yielded systematic differences in anchoring behavior.

Multiple Linear Regression. We fit an Ordinary Least Squares (OLS) regression model predicting the standardized anchoring effect, *AnchoringEffect_z*, from two categorical variables: *PresenceGroup* (HighPresence vs. LowPresence) and *AnchorGroup* (HighAnchor vs. LowAnchor). We also included their interaction to capture any synergistic effects. Table Table 1 reports the coefficient estimates, standard errors, t-values, and p-values.

Table 1: OLS Regression Results for Anchoring Effect

Term	Coef.	Std. Err.	t	p-value
Intercept	-0.1029	0.104	-0.991	0.325
C(PresenceGroup)[T.LowPresence]	0.0133	0.165	0.080	0.936
C(AnchorGroup)[T.LowAnchor]	-0.0031	0.147	-0.021	0.983
C(PresenceGroup)[T.LowPresence]:C(AnchorGroup)[T.LowAnchor]	0.1326	0.226	0.586	0.560

Model Statistics: $R^2 = 0.015$, Adjusted $R^2 = -0.024$, $F(3, 76) = 0.3886$, $p = 0.762$.

Table 2: Two-way ANOVA Results for Anchoring Effect

Source	Sum Sq	df	F	p-value
C(PresenceGroup)	0.1376	1	0.5550	0.4586
C(AnchorGroup)	0.0554	1	0.2234	0.6378
C(PresenceGroup):C(AnchorGroup)	0.0852	1	0.3436	0.5595
Residual	18.8448	76	–	–

A.3.2 Experiment 2. The following table shows the complete results of some verifications.

Table 3: Posthoc Wilcoxon Tests for Task Effect on Agency (Bonferroni-adjusted)

Comparison	W	p_{raw}	p_{adj}	**MedDiff**	**Direction**	**Sig.**
`Low_WithoutTask` vs `Low_WithTask`	118.50	0.3680	1.0000	−0.0833	No difference	n.s.
`Low_WithoutTask` vs `High_WithoutTask`	1.50	< 0.001	< 0.001	−1.3958	Higher in `High_WithoutTask`	*
`Low_WithoutTask` vs `High_WithTask`	0.00	< 0.001	< 0.001	−1.5833	Higher in `High_WithTask`	*
`Low_WithTask` vs `High_WithoutTask`	0.00	< 0.001	< 0.001	−1.3542	Higher in `High_WithoutTask`	*
`Low_WithTask` vs `High_WithTask`	3.00	< 0.001	< 0.001	−1.5000	Higher in `High_WithTask`	*
`High_WithoutTask` vs `High_WithTask`	86.50	0.1170	0.7020	−0.0417	No difference	n.s.

Table 4: Repeated Measures ANOVA for Confidence-Weighted Anchoring Effect

Source	**df**	**F**	**p**	η_p^2
Agency Level (High vs Low)	1, 23	0.0312	0.8613	< 0.001
Task (With Task vs Without Task)	1, 23	0.0516	0.8222	< 0.001
Agency Level × Task	1, 23	0.0505	0.8243	< 0.001

Table 5: Mean Confidence-Weighted AnchoringEffect_z by Condition

Condition	$\textbf{Mean}_{\text{weighted}}$
High WithTask/HighAnchor	−0.020
High WithTask/LowAnchor	−0.078
High WithoutTask/HighAnchor	−0.072
High WithoutTask/LowAnchor	−0.062
Low WithTask/HighAnchor	−0.009
Low WithTask/LowAnchor	−0.030
Low WithoutTask/HighAnchor	0.047
Low WithoutTask/LowAnchor	0.022

Table 6: Means and standard deviations of self-avatar overlap scores in each experimental condition.

Condition	Mean	SD
High With Task	4.13	1.32
High Without Task	4.18	1.36
Low With Task	3.60	1.30
Low Without Task	3.74	1.25

Table 7: ART-based ANOVA results examining the main effects and interaction of Agency and Task on Self-Avatar Overlap scores.

Source	SS	df	MS	F	p	η_g^2
Agency	29277.63	1	29277.63	5.03	.035	.019
Task	12364.69	1	12364.69	7.07	.014	.008
Agency × Task	63842.97	1	63842.97	35.25	$< .001$	.041