This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3549883

"One Body, But Four Hands": Exploring the Role of Virtual Hands in Virtual Co-embodiment

Jingjing Zhang, Xiyao Jin, Han Tu, Hai-Ning Liang, Zhuying Li, Xin Tong*

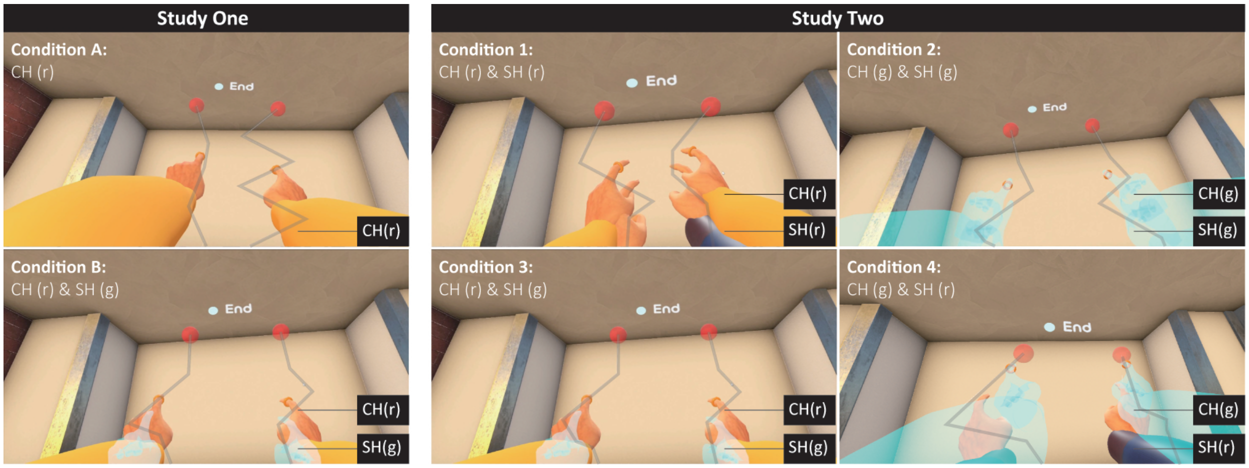

Fig. 1: Study One presented co-embodied hands alone (Condition A) and alongside users' self-hands (Condition B). Study Two further differentiated these experiences across four conditions: Conditions 1 and 2 showed identical hand transparency levels for co-embodied and user's self-hands, while Conditions 3 and 4 featured different transparency levels. *(Co-embodied hands & self-hands (CH/SH); ghost & realistic appearances (g/r)*

Abstract—Virtual co-embodiment in virtual reality (VR) allows two users to share an avatar, enabling skill transfer from teachers to learners and influencing their Sense of Ownership (SoO) and Sense of Agency (SoA). However, mismatches between actual movements and displayed actions in VR can impair user experience, posing challenges to learning effectiveness. Although previous studies have addressed the influence of virtual bodies' visual factors on SoO and SoA, the impact of co-embodied hands' appearances remains underexplored. We conducted two user studies to examine the effects of virtual self-hands' existence and their visual factors (transparency and congruency) on SoO, SoA, and social presence. Study One showed significant improvements in SoO and SoA with the existence of virtual self-hands. In Study Two, we kept the self-hands and further focused on hand transparency and congruency. We found that identical appearances between self-hands and co-embodied hands significantly enhanced SoO. These findings stressed the importance of visual factors for virtual hands, offering valuable insights for VR co-embodiment design.

Index Terms—Virtual Reality, Co-embodiment, Virtual Hand Representations, Body Ownership, Agency, Social Presence

1 Introduction

Skill acquisition can be facilitated through teachers guiding students with demonstrations, which promotes both skill learning and transfer [10, 48]. This method of instruction is particularly relevant for learning motor skills, including various practical activities related to sports [22] and industry [5]. Traditional motor skill learning is often based on learners observing the movements of a teacher from a third-person perspective in non-virtual reality (VR), which can lead to inaccuracy in novices' understanding of the movements and reduce learning efficiency [45]. Furthermore, prior work has demonstrated that allowing learners to observe teachers' movements from a first-person perspective (1PP) in VR helps them replicate these movements more accurately, enhancing immediate performance [28, 60]. Nevertheless, the use of such observations in 1PP has been shown to hinder long-term retention of motor skills, presenting challenges to sustained learning. Interestingly, a recent study found that using a virtual co-embodiment approach improved learning efficiency for motor skills related to declarative and procedural memory and promoted long-term retention of skills [53]. Another research also suggested that learners significantly improved their motor skill learning efficiency when they controlled co-embodied hands alongside an expert [35]. This improvement occurred because virtual co-embodied learning updated the learner's body schema, compared to observing teachers in 1PP or learning independently in VR [35]. Body schema, generally understood as an unconscious framework that facilitates motor actions, can be updated through body movements that are associated with a high sense of agency (SoA)[1] to assist achieve motor skills' long-term retention [6, 33, 35]. Specifically, learners need to experience strong SoA during exercises to support efficient skill transfer from experts [33]. However, there is a mismatch between the real hand movement performed by the learner and the corresponding co-embodied hand movement displayed in VR

[1] SoA refers to the subjective experience of control over movements and the belief that one is the owner of those movements [32].

- *Jingjing Zhang was a research assistant at the Hong Kong University of Science and Technology (Guangzhou) during this research period.*
- *Xiyao Jin is with University of Tsukuba, Japan.*
- *Han Tu was with the Massachusetts Institute of Technology, United States.*
- *Zhuying Li is with Southeast University, China.*
- *Hai-Ning Liang and Xin Tong are with the Hong Kong University of Science and Technology (Guangzhou), China.*

** Corresponding author: Xin Tong xint@hkust-gz.edu.cn*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

(controlled by the learners and another expert), which can negatively impact the experience of the learners.

SoA and sense of ownership (SoO)[2] are two main subcomponents of the embodiment, which describes how users represent themselves through an avatar in VR [19, 32]. Moreover, co-embodiment extends the embodiment concept by allowing two users or entities (e.g., another user, robot, or autonomous agent) to share a virtual avatar (the virtual co-embodied body) [8, 17, 34]. Body-part segmented and weighted approaches are the two main co-embodiment methods. The former assigns different parts of the body to each user for control [23, 24], and the latter allows two users to control the co-embodied body with differing degrees of influence [17, 33–35]. Previous research has explored how these two approaches using the co-embodied avatar influence the user experience for effective collaborative interaction [17, 34, 35]. One work using the body-part segmented approach found that participants' SoA and SoO in virtual co-embodied arms were noticeably higher when they had visual information on the shared goal or their partner's target, compared to the condition where the target was hidden [24]. Additionally, users could overestimate SoA in weighted-average co-embodiment when movements aligned and goals were shared [17]. Further research also explored how overestimating SoA enhanced learners' process of acquiring motor skills and strengthened their long-term retention, especially when using the co-embodied hand alongside teachers [33, 35]. Another research in virtual weighted-average co-embodiment explored the role of social presence[3] and its influence on the collaborative experience, and participants experienced interactive cues to enhance perceived social presence and aid movement coordination [58].

Studies on single-user embodiment showed that the visual appearances of virtual avatars (including the realism of their hands) significantly affected SoO by enhancing user perceptions of their virtual selves [2, 15, 27, 64]. Although these studies have shown that SoO decreases with less realistic and more transparent hands, they have not found significant effects on SoA [42, 43]. Meanwhile, recent research highlighted that self-avatar's congruence with the participant's own physical body yielded significant benefits for users' SoO and self-identification [40]. This finding suggested the potential influence of visual congruence on their SoO, although its effects on the SoA and social presence when they were co-located with multiple virtual humans of varying appearances remained unclear. In addition, the co-embodiment work has introduced an innovative idea involving a co-embodied avatar with a non-human appearance (i.e., a slug) [50]. However, there is a notable gap in research regarding how visual factors of virtual co-embodied hands affect SoO, SoA, and social presence when two users engage in collaborative interactions.

Current studies have underexplored whether the existence of virtual self-hands (users' own virtual hands) aside from the co-embodied hand impacts the interactive experiences during the weighted-average shared perceptual activity. Additionally, research has yet to investigate how visual factors of virtual hands (i.e., transparency and congruency) affect SoO, SoA, and social presence in the context of co-embodiment. In this work, we explore the following research questions (RQs):

RQ1: How do the existence of virtual self-hands beyond co-embodied hands affect users' SoO, SoA, and social presence in VR co-embodiment scenarios?

RQ2: How do the virtual hands' transparency (ghost or realistic) and congruency (identical or different) of virtual self-hands and co-embodied hands influence users' SoO, SoA, and social presence in the virtual co-embodiment?

In this paper, our research aims to explore the role of virtual hand designs under virtual co-embodiment through two user studies. User Study One examined how the existence of the user's own virtual hands affected users' SoO, SoA, and social presence. Findings revealed that the presence of virtual self-hands significantly improved users' SoO and SoA. In User Study Two, we retained virtual self-hands and investigated the effects of virtual hands' *Transparency* and *Congruency* between the self-hands and co-embodied hands on these three perceptions. The results showed that identical hand transparency improved users' SoO more effectively than the varying congruence condition. We asked participants to perform the bilateral coordination task in the co-embodiment, referring to the dual-task paradigm from previous VR co-embodiment work [35, 49, 53]. Additionally, we implemented a programmed virtual agent that collaborated with human participants on the movement task as their partner to precisely control the experiment and ensure consistent interaction dynamics, adopted from prior studies [8, 36, 54]. Overall, the contributions of this study are empirical insights into how virtual hands' transparency and congruency in co-embodiment scenarios impacted the SoO, SoA and social presence of users, listed as follows: (1) the existence of virtual self-hands enhances SoO and SoA in co-embodiment tasks; (2) aligning the congruent appearance of self-hands with co-embodied hands strengthens users' SoO, highlighting the role of virtual hand designs for interactive experiences; and (3) design implications for future development of self-hands and co-embodied hands in collaborative tasks, such as enhancing visual references through self-hands, reducing cognitive load with visual cues, and improving intention recognition.

2 Related Work

2.1 Co-embodiment in VR

Co-embodiment in VR is defined as a situation where two users share or control one virtual avatar, thus influencing its movements [17, 39]. This novel concept of virtual co-embodiment has been identified in recent literature as a scenario where an amalgamation influences the same virtual avatar or hand's movement, either *body-part-segmented based* or *weighted-average-based*, of multiple users' movements. The potential applications of co-embodiment could be monumental for collaborative tasks and skills learning, with studies suggesting the possibility of both users retaining the SoA and SoO over the co-embodied avatar [17].

Body-part-segmented based. An alternative approach to co-embodiment focuses on controlling distinct body parts, such as two users controlling the left or right limbs of the co-embodied avatar [23–25]. The work observed that the participants experienced stronger SoA and SoO for the virtual arm they controlled themselves compared to the one controlled by their partner [25]. The embodiment was increased for the partner-controlled arm when users shared a common goal or saw their partner's target. The visibility of the partner's intentions appears to enhance the embodiment in co-embodiment scenarios significantly [24, 25]. Another study indicates that a co-embodied avatar segmented by body parts offers a novel experience compared to traditional drum learning, although it is rated slightly lower in SoO and SoA [23].

Weighted-average-based. The weight-based approach to co-embodiment examines how different proportions of each user's movement influence the shared avatar, with 'weight' referring to this proportion [17, 20, 21, 33–35]. Early studies have explored user sensations towards the avatar based on five weight ratios (W0, W25, W50, W75, W100) [17, 33]. One work revealed that participants using a co-embodied avatar for half-weighted arm-reaching tasks reacted faster and moved with greater precision than they did when operating alone [20]. The findings of this work also suggested that sharing of body parts could modulate perceived SoO and SoA, with both increasing as the weight ratios increase [20]. Besides, methods to dynamically adjust weight and control transitional weight have been shown to more effectively maintain a robust SoA and enhance learning efficiency compared to static weight control, particularly in motor skill training [33, 34].

2.2 Skills Learning through VR Co-embodiment

Using co-embodied avatars in VR for skill learning is a transformative approach that enhances the effectiveness of physical movement training, including motor skills [31, 35, 53] and their applications [8, 41, 49]. The co-embodiment technique in VR enables learners to observe teachers' hand movements from a 1PP while simultaneously fostering a profound SoA over their own movements [53]. Moreover, a fixed target position facilitates the establishment of a robust SoA and SoO by making the

[2]SoO refers to the perceptual illusion that a person owns a virtual body or part of it [32].

[3]Sense of social presence refers to the feeling of being in the company of another sentient being, even if that presence is virtual [58].

teacher's movements more predictable for learners and increasing the learners' control weight of the co-embodied hands has been found to enhance their SoA and SoO [17].

Previous work has shown that virtual co-embodiment can improve participants' learning efficiency and long-term retention of motor skills that require declarative memory (i.e., memories that can build connections between specific instructions and movements) [53]. The learners observe and synchronize their movements with an expert using the co-embodied hand, allowing them to replicate the expert's movements more accurately in the long term with a strong SoA [20, 31, 35]. For example, studies have demonstrated that learning efficiency was higher when users practice with an expert through co-embodiment than when observing teacher movements from a 1PP alone or learning independently in VR [31, 35]. Research comparing the usability and behavior of co-embodied hands versus self-controlled solo and in VR found that co-embodied hands were easier to imitate, especially when the teacher's hands aligned with the learner's direction [31]. Moreover, users can overestimate their SoA when performing reaching movements using co-embodied hands, and it has been found that humans tend to give precedence to the movement patterns of co-embodied hands rather than their own movements during body sharing [20]. Another study emphasizes that accurately updating the body schema with high SoA during co-embodied training enhances the learners' ability to integrate observed movements with their own, thereby improving skill acquisition [35].

The application of co-embodied avatars also offers potential in skill learning contexts [8, 41, 49]. For instance, research into rhythm learning by non-musicians using co-embodied hands has revealed that this method maintains higher levels of SoA than traditional action observation [49]. In welding training, co-embodiment in VR can enhance collaboration between trainers and trainees, leading to more efficient skill transfer [41]. Moreover, integrating co-embodiment with movement observation therapy in rehabilitation settings allows virtual agents to effectively take on the role of therapists, which can transform patient engagement and improve therapy outcomes [8]. These applications underscore the potential of co-embodied avatars in enhancing educational and therapeutic experience in VR.

2.3 Sense of Embodiment and Social Presence

The sense of embodiment in VR encompasses two primary components: SoO and SoA [2, 32]. The SoO relates to the user's self-attribution to virtual avatars, while SoA concerns the experience of active movements of these virtual avatars. Studies have shown that coherent visual and haptic feedback from virtual limb movement can instill a powerful illusion of both SoO and SoA [2]. A striking correlation has been observed between the realism of virtual hand representations and the degree of body ownership reported. In fact, a more realistic or personalized virtual hand, resembling the user's actual hand, has been shown to improve SoO [27, 32]. Interestingly, there is a debate in the literature on the impact of visual factors related to the appearance of virtual bodies on the SoA. More realistic virtual avatar representations have been reported to be associated with a stronger SoA [27, 44] or a lower SoA [2], while others have not found significant differences [37, 64]. The relationship between visuomotor congruence and SoA would be more pronounced than the visual stimuli provided by the appearance design of the virtual body. Studies corroborated that an alignment between actual and virtual movements substantially strengthens the individual's sense of control over their movements, and the SoA is likely to decrease when visual and motor signals do not align [16, 52].

Recent VR and AR research in collaboration reveals various effects of virtual avatars on social presence. The research found that avatars created with movement capture technology enhanced the sense of presence and achieved higher co-presence by evaluating avatar completeness and movement implementation [26]. Cartoon-like avatars were also reported to increase co-presence, while realistic avatars increased participant trust [29]. Other studies show that the visibility of different parts of the avatar body was found to influence the social presence experienced by local participants. In contrast, neither realistic nor low polygon avatars significantly impact social presence during tasks [61]. Similar contradictory results proposed that there was no significant effect of realistic humanoid avatars on co-presence [15], and avatar facial characteristics have no impact on social presence [13]. The effect of three levels of avatar transparency for task-centered collaboration on social presence was evaluated in a recent work, which indicated that highly visible avatars are not essential in scenarios where the primary goal is the completion of tasks rather than social interaction [62]. These contrasting findings indicate a nuanced interplay between avatar appearance and its impact on user interaction, calling for further investigation into how different avatar features affect virtual collaboration.

2.4 Visual Factors in Co-embodiment

While prior work on user interaction in VR has emphasized the pivotal role of visual factors, limited research has investigated the influence of co-embodied' visual factors. An article reported on an approach to enhance SoO, SoA, and learning performance between users by harnessing users' visuomotor training in VR [35]. Two users complete interactive tasks by controlling a co-embodied body, where the user can follow the task by sharing the perspective of another user (as a teacher). Although visually, it appears that four hands are simultaneously operating, with one pair representing the shared hands and another viewed through the shared perspective, the investigation into the visual cues that signal the existence of self-hands remains unclear. This area of study involves an overlap where one pair of self-hands visually coincides with the co-embodied hands. Additionally, only one study has explored the concept of a shared avatar with a non-human body appearance. This preliminary exploration seeks to expand our understanding of how the design of virtual bodies' appearances influences users' perceptions of co-embodied interactions, although it does not specifically measure perceptions related to SoO, SoA, and social presence [50]. In contrast, another study in a general shared virtual space investigates the impact of appearance congruence between one's self-avatar and other users on SoO. The findings indicate that appearance congruence enhances SoO and self-identification, while incongruence reduces these perceptions and affects the immersion experience [40].

In the context of embodiment, research has explored several visual factors, including the realism or transparency level of the avatar [1, 34, 42]. The visual portrayal of virtual hands can range from basic, stylized shapes to detailed, realistic renderings. Studies have shown that the level of detail in the appearance of users' hands can affect user reactions within VEs [2, 37, 64]. The findings highlight that users can follow tasks more easily with transparent avatars and that a teacher's perspective can amplify learning efficiency [56]. Transparent avatars that offer body imagery have been suggested to effectively convey body cues without dominating the visual field. They also experience less occlusion compared to other avatars of different fidelity, making them ideal for collaborative settings where user presence is necessary but should not obstruct the task [62]. However, the influence of visual factors in representing movement behaviors in co-embodiment remains an area with limited understanding.

3 The Design and Implementation of the Collaborative Drawing Task

We implemented a collaborative VR drawing task (bilateral coordination task) for two studies where participants used both hands to draw trajectories simultaneously in collaboration with a virtual agent. Moreover, we opted to incorporate a virtual agent into our studies to improve experimental control, allowing for precise manipulation of experimental variables and ensuring consistency for our participants in the studies. The virtual agent acts as a realistic, responsive partner programmed to emulate human-like interactions, and it manages the dual-hand task and serves as a substitute for a real human partner in this experiment.

3.1 Apparatus

We implemented the interaction techniques and experiment tasks in VR. The experiment is conducted using a Meta Quest 2 headset with its controller, allowing users to adjust the interpupillary distance for comfort. The experiment runs on the open-source Godot 4.0.2 game

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3549883

engine, utilizing its XR toolkit and keeping the FPS above 90. Participants would interact through an avatar featuring an upper body with hands and arms, where movements are driven by inverse kinematics. The VE is a default room with basic elements of the wall and floor, providing a simple and controlled space for testing. During the experiment, participants would use two controllers to complete the drawing tasks, with the controllers in the virtual world being replaced by virtual hands. We ensured that the participants could complete the tasks sitting comfortably and did not collide with other objects in the real world.

3.2 The Co-Embodiment Approach

For co-embodiment, we use the weight adjustment method [33, 34] to let the participants share the position and rotation control with the programmed virtual agent. The participant's control weight of the 'co-embodied hands' is 50%. The position and rotation of the "co-embodied hands" can be calculated based on the weighted average of the controller's and the agent hand's transform. The following equation expresses the calculation process.

$$T_c = W * T_s + (1 - W) * T_a \tag{1}$$

W Participant's control weight [0,1]. T_c Transform of the co-embodied hand. T_s Transform of self-hands. T_a Transform of the agent's hand.

3.3 The Task Design

The use of two hands performing a task has been explored in several previous co-embodiment studies on motor skill learning, the right hand drawing a seven-pointed star and the left hand drawing a five-pointed star simultaneously [35, 54], beating separate drums with both hands [49], or using both hands to draw the virtual star together [53]. This dual-task paradigm has been reported to require one user to perform two relatively simple motor tasks simultaneously and is a commonly used method to assess motor skill learning [30, 46, 57]. Meanwhile, bilateral coordination tasks in psychology are one type of dual task, which involves the use of both hands for distinct but coordinated movements [18]. Building on this framework, our work designed a bilateral coordination task in which participants were required to draw distinct trajectories simultaneously with both hands (Fig. 2a). The user could cooperate with a virtual expert agent to draw the lines between the starting points and the endpoints. Although the start and end points of each trajectory are fixed, the intermediate trajectory is generated by a path generation algorithm. For each participant, the trajectories that appeared in each task were different, avoiding the influence of the learning effect during the experiment and preventing participants from becoming familiar with similar patterns. Moreover, data on trajectories and participant drawings were collected during the experimental procedure.

Participants were asked to use co-embodied hands and self-hands to complete two experiments, and these virtual hands with two transparency levels (realistic or ghost hands) were designed in our studies. In previous studies, realistic virtual hand designs featured similar human textures with no transparency (i.e., a skin color similar to all participants), while ghost virtual hands with high transparency levels were used to distinguish the additional parts from the rest of the body (see Fig. 1). For these virtual hand representations, we use different shader programs to change the virtual hands' material and visibility. Specifically, realistic hands use an opaque shader with no transparency level (realistic appearance), and ghost hands use a transparent shader with a purely blue color, with a high transparency level of 70% (ghost appearance). We distinguished between two types of virtual hands to reduce users' visual cognitive load in VEs: co-embodied hands and self-hands. We marked the co-embodied hand with a ring on the index finger and rendered its arm yellow. In contrast, users' self-hands are rendered in dark blue without jewelry (Fig. 2).

3.4 The Virtual Agent

It has been reported that the expert in the co-embodiment is not necessarily a real human, as teaching requires both high professional skills and the ability to remain unaffected by the learner's movements [8, 49, 54].

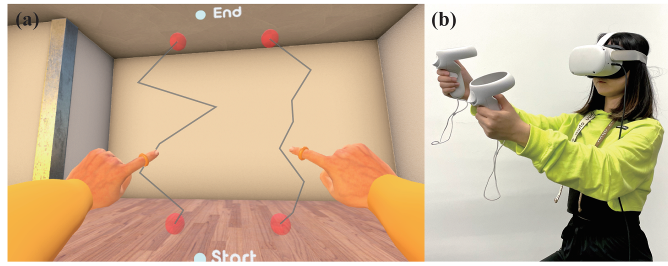

Fig. 2: (a): The co-embodied hands were used in VEs for participants to draw two distinct trajectories simultaneously (Condition A without the existence of self-hands in User Study one as an example); (b) Participants owned virtual hands to perform the drawing task.

James and Michael [49] presented a drumming learning approach that the user co-embodied with a programmed avatar, and Zhang and Yu [8] also suggested the robot guidance method instead of therapists. Moreover, similar idea had also been pushed forward to co-embodiment with an AI teacher, whose movement was trained in human teaching behavior [54]. In our study, participants (i.e., as novices) collaborated with a programmed virtual agent (i.e., acting as an expert) on the bilateral coordination task, and were able to move their virtual hands synchronously to complete the 'Drawing Phase' (explained in Section 3.5). The moving speed of the agent's hands was not constant, and we programmed the speed with the *'EaseInOutQuad'* easing function (with the speed interpolation slowest at both ends) to make the movement not too rigid and easier to follow by the participants (Fig. 3d). The equation of this function below explains the position change by time 't'. The total time for the agent to finish drawing is fixed, it accelerates in the first half ($t < 0.5$) and decelerates in the second half ($t > 0.5$), creating a natural easing effect.

$$\text{Position}(t) = \begin{cases} 2t^2, & \text{if } t < 0.5, \\ 1 - \frac{(-2t+2)^2}{2}, & \text{if } t \geq 0.5. \end{cases} \tag{2}$$

The virtual agent's movement was fully constrained by the 'PathFollow3D[4]' function integrated by Godot 4 game engine, which ensured the accurate performance of the agent's drawing trajectories. Considering the teaching purpose, the expert (i.e., the virtual agent) was supposed to lead novice users in learning motor skills, and we set the virtual agent with high accuracy in completing the task.

3.5 The Three Task Phases

In our experiment, participants are required to complete similar dual-hand drawing tasks under various conditions, each of which presents different visual combinations of virtual hands. One trial of each task included three phases (Fig. 3): *Prepare phase*, *Drawing phase*, *Complete phase*. In the *Prepare phase* (Fig. 3a), participants only saw a pair of abstract hands, indicated by spheres that mark the positions of their physical hands. After placing the position indicators in the 'start point' places and holding the grasp triggers on the controllers, they would see a three-second countdown. Then, the abstract hands were transformed into our co-embodied hands (or ghost appearance) in the experiment. Participants experienced co-embodiment with the virtual agent to complete the drawing task using both hands during the *Drawing phase* (Fig. 3b). The participants would see different visual combinations of virtual hands under various conditions. Finally, when the co-embodied hands reach the 'endpoints', the *Complete phase* (Fig. 3c) could be triggered, and after 2 seconds, the participants' virtual hands would change back to the abstract appearance. After one trial ends, it automatically progresses to the next. The participants were required to complete six trials on each condition.

[4] https://docs.godotengine.org/en/stable/classes/class_pathfollow3d.html

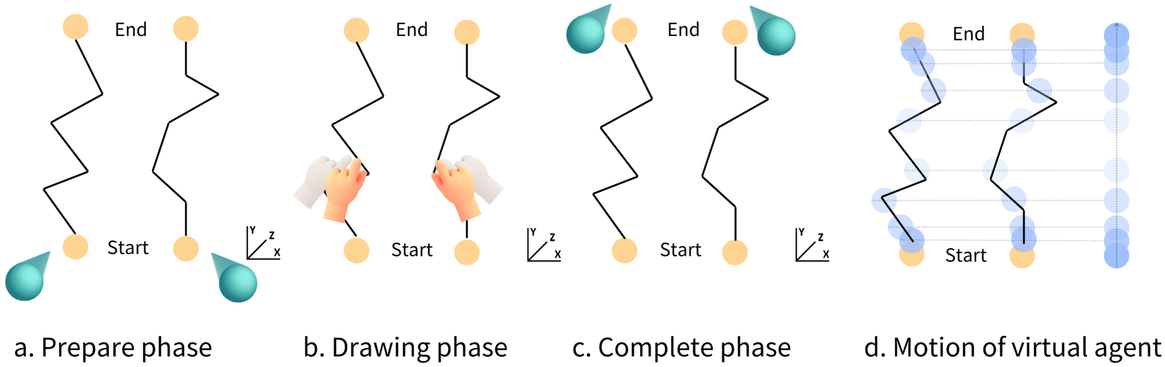

Fig. 3: The three phases of the collaborative task adopted in User Study One and User Study Two: In the prepare phase (a), the user's hands are represented by abstract geometry; in the drawing phase (b), the user needs to complete the drawing task; after that (c), the user's hands will return to the abstract form. The blue circles in (d) show the movement of the virtual agent during the drawing task by 'easeInOutQuad' function. The virtual agent acts as the expert in the bilateral coordination task.

4 User Study One

After the approval of the Institutional Review Board (IRB) at Xi'an Jiaotong-Liverpool University (approval number ER-LRR-12889407202309 28115806), we conducted User Study One with two different visual conditions to understand how visual factors of users' hands (whether self-hands exist or not) influence SoO, SoA, and social presence in co-embodiment tasks. Condition A only displayed realistic co-embodied hands, and Condition B displayed co-embodied hands and self-hands with realistic and ghost appearance, respectively.

4.1 User Study One Method

4.1.1 Participants

A total of 15 participants (3 females and 12 males, aged 23.80±2.93) joined this experiment. All participants self-reported themselves as novice users of VR co-embodiment scenarios. We offer a gift equivalent to USD 10 in the local currency to compensate for their time participating in the study.

4.1.2 Study Design

The independent variable of User Study One was *user's self-hands (SH) is displayed or not*. There are two levels in this experiment: co-embodied hands (*CH-r*) in realistic appearance or dual hands including co-embodied realistic hands and self-hands in ghost appearance (*CH-r* and *SH-g*), which were used to explore the influence on dependent variables (SoO, SoA, and social presence).

Condition A: CH(r) In this condition, participants could control realistic co-embodied hands with the virtual agent to collaboratively engage in the task of drawing lines, each contributing equally to the control. Specifically, the control was split evenly, with both the participant and the avatar responsible for 50% of the movements.

Condition B: CH(r)& SH(g) The primary difference from Condition A is the addition of virtual self-hands (i.e., participants' own hands with ghost appearances). Participants could simultaneously observe the movements of both realistic co-embodied hands and their self-hands in the line-drawing task. Notably, the ghost self-hands are entirely controlled by the participant, highlighting their direct involvement in the task. The co-embodied hand in this condition, the same as in Condition A, represents a weighted combination of the movements from both the participant and the collaborative virtual agent involved in this study.

4.1.3 Procedures

Before the experiment, the participants were briefed on tasks and objectives and provided an overview of the testing procedure (Fig. 3). We informed participants that they would be completing a collaborative task with another individual in a separate room, requiring them to control the co-embodied hand. Before the experiment, the context was set without revealing that the experimental manipulation involved an agent (i.e., not a human collaborator), unbeknownst to participants. Their understanding was ensured before signing an informed consent form. Participants were also informed to stop the experiment anytime if they felt uncomfortable and then allowed to use the VR Head-Mounted Display (HMD) and its controllers to start the experiment. To ensure the accuracy and reliability of the testing, all equipment underwent thorough debugging and calibration before the experiment commenced. For example, calibration included adjusting the interpupillary distance settings on HMD to match each participant's unique measurements, ensuring a natural and correct viewpoint in VEs.

Tutorial: Participants were familiar with using self-hands for drawing tasks in VEs, and the experimenter explained the background of the cooperative task and the experimental considerations. After entering the virtual scene, the participants were presented with a simple explanatory scenario that illustrated the main form of the testing task, allowing them to become accustomed to the overall task procedure. The experimenter guided the participants in understanding the basic requirements of this task, and they were allowed to practice plotting a trajectory between two points (Fig. 2).

Main Tasks: Participants were randomly assigned to one of two testing conditions after the tutorial, and they completed the experiments in a random order. Participants collaborated with another user (i.e., a virtual agent in this study) to control co-embodied hands in VEs. They completed a dual task in which their left and right hands depicted different trajectories, each with a predetermined start and end point set along 3D paths. The co-embodied hand's movement was calculated using a weighted average of the participant's and virtual agent's position and speed (50% weight control each). Each task was considered complete when the participant successfully guided the co-embodied hand to the endpoint. Each participant repeated the main tasks in six trials under each condition. It should be noted that to ensure a better collaborative experience for the participants, the experimenter gave them an oral command to press the VR controller button before each trial. This oral prompt was crucial for synchronizing the start of the activity, helping the participants understand the method of simultaneous cooperation between the two individuals. Once the button is pressed, the combination of virtual hand representations used in our experiment appears. During the same time, participants could see a 3-second countdown appearing in their view. They needed to wait for the countdown to finish and then start moving their virtual hands, following the trajectory instructions.

The behavior performance of the participants during the main tasks was recorded to evaluate the quality of their interaction. After completing the tasks under each condition, the participants were allowed to remove the headset for a short break. They needed to answer a 7-point Likert scale questionnaire on SoO, SoA, and social presence on the computer web page. They were given a 3-minute rest period between the two conditions to alleviate fatigue. Excluding the tutorial, each experimental condition (including the main tasks and questionnaire) took approximately 5 minutes to complete. In total, with breaks, the entire session lasted about 15 minutes.

4.1.4 Measurements

We adopted the within-subject design for the experiment and also surveyed the influence of the following measures to assess the effect of users' SoO, SoA, and social presence. Detailed questions can be found in Table 1, and the questionnaire was shown simultaneously in English (original) and the participant's native language (translated).

SoO. The ownership questionnaire in Table 1 primarily measures and understands users' cognitive perception of virtual co-embodied hands and seeks to quantify the extent to which individuals perceive their virtual hands as their own within a given context [17,33,47].

SoA. The agency questionnaire in Table 1 fundamentally aims to comprehend users' perception of controlling their movements by co-embodied hands in VR and strives to quantify the degree to which individuals perceive their own movements as self-generated [17,33,47].

Social Presence. The questionnaire in Table 1 assesses individuals' perception of the presence of others in VEs during the co-embodied hands' interaction, quantifying the extent to which users perceive a sense of shared presence with others within a sane VR context [3,7,15].

Table 1: Detailed questions of user perceptions

SoO (Q1-Q5); SoA (Q6-Q11); Social Presence (Q12-Q15)
Q1.I felt as if I was looking at my own hand.
Q2.I felt as if the virtual co-embodied hand was part of my body.
Q3.I felt as if the virtual co-embodied hand was my hand.
Q4.I felt as if I had no longer a right hand, as if my hand had disappeared.
Q5.I felt as if the virtual co-embodied hand was from someone else's body.
Q6.The virtual co-embodied hand moved just like I wanted to, as if it was obeying my will.
Q7. I felt as if I was controlling the movement of the virtual co-embodied hand.
Q8.I felt as if the virtual co-embodied hand was controlling my will.
Q9.I felt as if the virtual co-embodied hand was controlling my movements.
Q10.I felt as if the virtual co-embodied hand had a will of its own.
Q11.I felt as if I was causing the movement of the virtual co-embodied hand I saw.
Q12.To what extent did you feel the other player was with you in VEs?
Q13.To what extent did you have a sense of the emergence of a team during these events?
Q14.To what extent did you have a feeling that you were collaborating with real people and not robots?
Q15.To what extent did you have a sense of being 'part of the team'?

4.2 User Study One Results

The IBM SPSS statistical analysis software (Version 26) was used to perform all data in User Study One. We assessed the normality of all self-reported ratings to evaluate the scores of SoO, SoA, and social presence first using the *Shapiro-Wilk test*. All three data were found to follow normal distributions ($p > 0.05$), confirming the suitability of using parametric tests. Subsequently, we performed the paired t-tests to compare the two conditions among these three data. Fig. 4 presents the results of user perceptions from User Study One.

SoO. The t-test revealed a significant difference in the effect of owning virtual co-embodied hands between the two conditions ($p < 0.05$). The analysis showed a lower mean SoO score (M=3.53; SD=1.19) in Condition A (i.e., only co-embodied hands with realistic appearances). In contrast, the co-embodied hand in Condition B, which involved both the shared and self-hands, received a significantly higher SoO rating (M=4.32; SD=0.79). The lower score suggests a weaker connection between the participant and the co-embodied hand. It is theorized that SoO arises when there is a congruence between sensory cues, such as visual and proprioceptive feedback, and the mental representation of one's own body. The observed increase in SoO in Condition B could be attributed to the existence of self-hands, which could enhance the embodied sensation of the participants, reinforcing their perception of the virtual hands they see in VEs as being part of their own body.

SoA. A significant difference in SoA was observed between the two conditions ($p < 0.05$). Condition B featuring the self-hands reported a higher mean agency value (M=4.35; SD=0.97) compared to the lonely co-embodied hand condition (M=3.45; SD=0.94). The SoA refers to the subjective experience of control over movements and the belief that one is the owner of those movements. A heightened SoA is typically associated with the ability to predict the consequences of one's movements and a strong feeling of control over them. The existence of self-hands in VEs likely provided greater congruence between intended movements and visual feedback, thereby enhancing the participant's sense of control.

Social Presence. Statistical analysis revealed no significant difference between the two conditions on the sense of social presence ($p > 0.05$). Condition A with a single co-embodied hand revealed a lower mean value (M=4.12; SD=1.37) than other situations (M=3.98; SD=1.57). The sense of social presence refers to the feeling of being in the company of another sentient being, even if that presence is virtual. In this study, similar levels of social presence reported in both conditions suggest that the mere representation of a hand, whether it is a co-embodied hand or another form, does not significantly influence the participants' feelings of social presence. This could imply that other factors might play a more crucial role in fostering a sense of social presence.

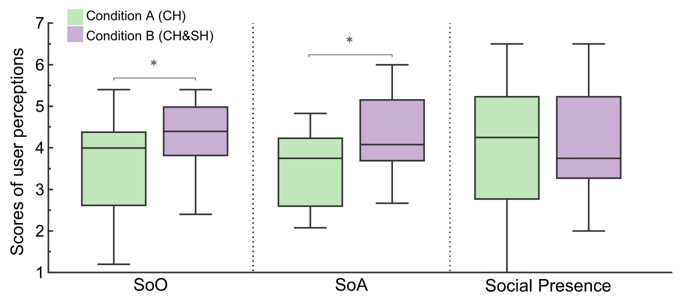

Fig. 4: The scores of user perceptions under two conditions in User Study One (CH: co-embodied hands; SH: self-hands).

5 User Study Two

Results from Study One indicated that the existence of self-hands in VR co-embodiment tasks impacts the collaborative experience of users. The visualization of the participants' own hands with a ghost appearance appeared to have a positive effect. Therefore, User Study Two further explored other settings of the virtual hand representations with two transparency levels. This study also received IRB approval from the university (same as before), and we prepared four different visual settings and allowed participants to take part in four groups of tasks randomly. Both settings included two pairs of virtual hands (co-embodied hands and self-hands) that appear in the virtual space. Two factors would be: '**Hand Transparency** with realistic and ghost appearances' and '**Hand Congruency** with same appearances on these two hands or not'.

5.1 User Study Two Method

5.1.1 Participants

A total of 21 participants (16 identified as males and five as females) from the university community joined the tasks. Their ages are between 19 and 31 years (M = 23.52±2.67 (SD)). All of them had normal or corrected-to-normal vision and were self-reported novices of VR co-embodiment, with 4 having no prior VR experience. They were also compensated as in Study One.

5.1.2 Study Design

In this study, we set up four combinations of co-embodied hands and self-hands for participants to explore which types of visual settings would work better. Before the experiment began, the experimenter explained the background and purpose of the experiment to all participants. This experimental procedure in User Study Two would be consistent with that of Study One, with the addition of different conditions. Upon entering VEs, participants were first familiarized with a tutorial task to understand the task requirements and procedural flow. To ensure participants' cooperation synchronicity, all were required to initiate the cooperative task according to the researcher's verbal initiation instructions. Under the researcher's guidance, they practiced plotting trajectories between two points in the dual-task scenario. After the tutorial, participants were randomly allocated testing scenarios, each aiming to guide co-embodied virtual hands to a trajectory endpoint. A three-minute intermission was provided between conditions. Upon completion of the condition, participants assessed their perceptions of SoO, SoA, and social presence over the co-embodied hand via the 7-point Likert scale questionnaire. The items in this questionnaire were consistent with those used in Study One, as detailed in Table 1. At the end of all the experimental tasks, we conducted semi-structured interviews (see Table 2 for details) to obtain additional information on the experiences and perceptions of the participants.

Based on dual virtual hands combination, there are two dimensions of independent variables here: ***Hand Transparency*** (two levels: ghost (g) and realistic (r) hands) and ***Hand Congruency*** (two levels: identical appearances of *CH & SH* and different appearances of *CH & SH*). The dependent variables included SoO, SoA, and social presence. Four different conditions are explained below, see Fig. 1.

Condition 1: CH(r) & SH(r) Two identical virtual hands with realistic representations would appear, aiming to facilitate participants in performing the drawing task with both hands.

Condition 2: CH(g) & SH(g) To perform the drawing task with both hands in the virtual world, two identical virtual hand appearances have appeared. Both the participant's self-hands and the co-embodied hands were depicted with a ghost representation.

Condition 3: CH(r) & SH(g) Two different virtual hand representations of *co-embodied hands and ghost hands* appeared in VR. Contrary to Condition 1, the self-hands were depicted with a ghost appearance, while the co-embodied hand was shown with a realistic appearance.

Condition 4: CH(g) & SH(r) Similar to Condition 3, participants performed the drawing task using two different Hand Transparency. Self-hands with realistic appearances are of the participant's own, while co-embodied hands are represented with a ghost appearance.

5.1.3 Measurements

The measurements used to evaluate users' SoO, SoA, and social presence in User Study Two were similar to those used in User Study One. Meanwhile, User Study Two incorporated the collection of interview data and details of the interview questionnaire can be found in Table 2.

Table 2: Detailed interview questions

1. How did you feel interacting with the virtual hands you saw? How does their appearance, both your virtual self-hands and the co-embodied hands, influence your interaction?
2. What are your feelings towards the appearance of these virtual hands, both co-embodied hands and self-hands? How does this impact your usage experience and process?
3. Which combination of virtual hand appearances do you prefer? Why?
4. Can you list some advantages and disadvantages of these virtual hands' appearance with co-embodied hands and self-hands?
5. Did you experience any discomfort or trouble using different virtual hand designs? Can you explain further?
6. If you could change the appearance of these virtual hands (i.e., co-embodied hands and self-hands), what would you like it to look like?
7. Did you have any other notable experiences while using them? Were they related to virtual hands' appearance?

5.2 User Study Two Results

We aimed to evaluate the effects of '*Transparency*' and '*Congruency*' on these dependent variables across four conditions. The factors analyzed were **Hand Transparency** (levels: g and r) and **Hand Congruency** (levels: *identical* and *different*), both treated as within-subjects factors. The SoO, SoA, and social presence according to users' scoring guidelines were shown in Fig. 5, and the descriptive statistics of these dependent variables were evaluated under each condition.

The same statistical analysis software was used to perform the primary data analysis in this study, and the normality of data was first assessed with the *Shapiro-Wilk tests*. The scores of SoO and social presence met the normality assumption ($p > 0.05$) and were analyzed using two-way repeated measures ANOVA. If the assumption of sphericity by *Mauchly's test* was violated, the degrees of freedom were corrected using *Greenhouse-Geisser estimates*. Then, the non-normal data for SoA ($p < 0.05$) was evaluated using the *Shapiro-Wilk test*, and the *ARTool* library of R software (version 4.3.1) was used to assist with the Aligned Rank Transformation (ART) method before further analysis [59].

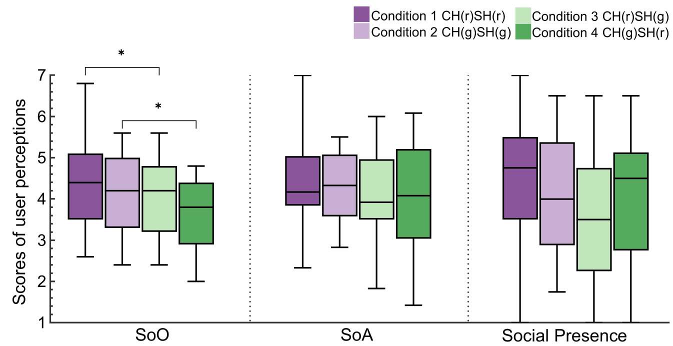

Fig. 5: The scores of SoO, SoA, and Social presence under four conditions from Condition 1 to Condition 4 in User Study Two.

SoO. The two-way repeated measure ANOVA was conducted for SoO to evaluate the effects of *Hand Transparency* and *Hand Congruency*. The main effect of transparency was not significant, F(1, 20)=0.180, $p > 0.05$, indicating that there was no substantial difference between ghost and realistic representations. The main effect of congruency was also not significant, F(1, 20)=0.680, $p > 0.05$, suggesting that the co-embodied hands and self-hands did not differ significantly in their effect on SoO. However, a significant interaction between *Hand Transparency* and *Congruency* was observed, F(1, 20)=0.032, $p < 0.05$, $\eta^2 = 0.210$, indicating that the effect of one independent variable on SoO depends on the level of the other variable. *Post-hoc* analyses using *Tukey's HSD* were conducted to investigate the nature of this interaction further.

The analysis revealed that within the co-embodied hand with ghost appearances in Condition 4 (CH(g) and SH(r); M=3.71; SD=0.80), and Condition 2 (CH(g) and SH(g); M=4.19; SD=0.98), the difference between the ghost and realistic hand was significant ($p < 0.05$). Meanwhile, there was also a significant difference between Condition 3 (CH(r) and SH(g); M=4.02; SD=1.02) and Condition 1 (CH(r) and SH(r)); M=4.36; SD=1.05; ($p < 0.05$). This suggests that when two virtual hands (i.e., co-embodied hands and self-hands) were present at the same time, the transparency differences between the two hands significantly affected users' ability to distinguish between these two virtual hands. The lack of significant main effects for independent variables on owning the virtual avatar suggests that each independent variable alone does not influence the dependent measure. However, the significant interaction effect highlights the complexity of these variables' relationship, demonstrating that the impact of hand transparency is contingent upon the level of the virtual hands.

SoA. The data for SoA were first processed using the ART procedure in the R software, which aligns and rank-transforms the data, to address the violation of the normality assumptions. ART-based ANOVA was then conducted to examine the main effects and their interaction. Investigation of the effects of *Hand Transparency* and *Hand Congruency* on SoA yielded the following insights: transparency's influence did not reach statistical significance, F(1, 20)=0.027, $p > 0.05$, suggesting that this variation had no discernible impact on SoA. The similar pattern emerged for the congruency, without a significant effect detected, F(1, 20)=0.061, $p > 0.05$. Furthermore, the expected combined influence of the independent variable, as evidenced by their interaction term, was not statistically supported, F(1, 20)=1.785, $p > 0.05$. This outcome

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3549883

suggests a lack of synergistic effect between them in relation to the SoA. Although the differences between these data of SoA appear to be slight, the M and SD for each of the four conditions were as follows from Condition 1 to 4: CH(r) & SH(r) (M=4.48, SD=1.20), CH(g) & SH(g) (M=4.29, SD=0.89), CH(r) & SH(g) (M=4.04, SD=1.10), and CH(g) & SH(r) (M=4.16, SD=1.26).

Social Presence. The two-way repeated measures of ANOVA were performed separately for the sense of social presence. The results indicated that there are no significant main effects of *Hand Transparency* (all $p > 0.05$) and no significant main effects of *Hand Congruency* (all $p > 0.05$) for any of the dependent measures. Furthermore, none of the interaction effects between them reached statistical significance (all $p > 0.05$) for these variables. While the social presence was significantly rated higher values in Condition 1 (CH(r) and SH(r)); M=4.37; SD=1.58) over Condition 3 (CH(r) and SH(g); M=3.64; SD=1.62; $p < 0.05$).There are the other detailed descriptive statistics: Condition 2 CH(g) & SH(g) (M=4.10, SD=1.57) and Condition 4 CH(g) & SH(r) (M=3.94, SD=1.56).

6 Discussion

6.1 Perceptional Influence in VR Co-embodiment Task by Visual Cues on Virtual Hands

Significance of Self-Hand Existence in Enhancing SoO and SoA. User Study One initially compared two distinct scenarios: both self-hands and co-embodied hands and only co-embodied hands. These conditions highlighted the crucial role of self-hands in enhancing the SoO within a co-embodied environment. When participants observed virtual hands that mirrored their own, they were more likely to perceive these hands as extensions of themselves. This perception did not apply in the cases where the hands were shared. Strengthening the SoO, as demonstrated in previous research, enhanced their sense of belonging to the virtual world [2, 17, 27, 33]. Such findings underscore the importance of the visibility of one's own body parts, echoing the results of previous research that has similarly highlighted this effect [38]. Another study also revealed that several participants preferred having a dual representation to more effectively perceive the position of their real arm [12].

The existence of self-hand also correlates with increased SoA, which could reflect the users' heightened perception of influence and control in VEs where their movements are visible and manipulable directly. It is reasonable to assume that providing users with full control over their hands (100% weight control for self-hand) in Condition B (CH&SH) would enhance the SoA. Participants used a pair of co-embodied hands with limited control (50% weight control for the co-embodied hand), a reduction in SoA can be observed in Condition A (CH). This finding is consistent with previous research [17, 33], which has consistently shown that the degree of control over visual feedback of a person's movements in a virtual space affects users' SoA. Moreover, a recent study explored the impact of users' own hands with two distinct representations: one representing the real hand movements and another with distorted movements designed to perform remote interaction tasks using the manipulation technique [12]. The findings revealed that the dual visual appearances of the virtual hands (including two hands with distorted and real movements) induced a higher SoA compared to when only real-hand movements were displayed [12]. Current evidence indicates that when users perceive virtual hands that move and act in alignment with their physical hands, they report a stronger sense of embodiment, including SoO and SoA [16, 17, 27, 52].

Impact of Visual Factors: *Hand Transparency* and *Hand Congruency* Design. In User Study Two, we analyzed the effects of 'ghost hands' with high transparency versus 'realistic hands' without transparency for two conditions with different hand appearances. Although transparency level alone did not significantly influence SoO, the interaction between different transparency levels underscored a complex dynamic in how visual factors affect user experience in VR. This result challenges the conventional expectation that higher visual realism enhances ownership feelings and suggests that differences in visual transparency may affect user engagement and perception [2, 37]. Interestingly, a previous study under dual hand representations reported that SoO and SoA scores were higher with realistic appearances than with ghost hands, highlighting the potential impact of hand factors on the embodiment [12]. In our conditions with different hand appearances (Condition 3 CH(r)&SH(g) and Condition 4 CH(g)&SH(r)), the combinations of virtual hand representations on the co-embodied hand and the self-hand had a clear distinction: one with a high level of transparency as 'ghost hands' and another with no transparency as 'realistic hands'. Although the independent variables on their own did not significantly sway the SoO, their interaction revealed a pivotal impact, underscoring the complex nature of virtual co-embodiment. The absence of significant main effects has shown that higher realism uniformly enhances ownership sensations and instead directs attention to special visual representations in interactive experiences [2, 27, 64].

The findings of User Study Two also explored the effect of hand congruency between self-hands and co-embodied hands. When participants were presented with dual representations of hands in virtual scenarios that maintained identical appearances, they reported higher SoO compared to other conditions that feature mixed representations, such as in Condition 1 (CH(r)&SH(r)) and Condition 2 (CH(g)&SH(g)). This observation highlights the importance of visual and operational consistency in reducing cognitive load, as participants may not have to spend time distinguishing between two sets of hands while performing tasks. Furthermore, this consistent visual theme allows users to integrate their movements and perceptions more easily within the VE, enhancing both the SoO and the overall immersive experience. This interaction suggests that the difference in appearance between co-embodied hands and self-hands could indeed elevate user perception, and this effect emerges only when both hands in identical appearance are visually juxtaposed within the virtual scenario. A previous study reported that having two representations helped users better understand their virtual movements and showed a good level of acceptance for the ghost hand in representing real movements [12]. Additionally, when users are asked to rate their SoO over the co-embodied hands, the difference observed may also stem from the fact that they do not require additional cognitive load or perceptual effort to differentiate between the two sets of hands.

The SoA observed in User Study Two is largely consistent with previous research [37, 64], indicating that the visual appearance of virtual hands (including hand transparency and hand congruency) has a relatively minimal impact on SoA. SoA is defined as the user's perception of control over their movements and the resulting consequences in VEs, which is governed primarily by factors related to the user's interaction with the system [32]. Factors such as the immediacy and accuracy of feedback to the user's movements and the consistency of interaction may have a more direct impact on SoA than the visual fidelity of the avatar's hands. Another explanation for this is that users are likely to experience a strong SoA primarily due to the clear causal relationship between their intended movements and the resulting outcomes within the VEs, which is influenced by the responsiveness of the situation. For instance, the social presence or the sense of being with another entity might be influenced more by the responsiveness and behavioral cues of the virtual others than by their visual realism [15, 61]. This sense of social interaction and connection could be affected by the quality of communication, shared goals, and collaborative tasks [13, 62].

6.2 Visual Guidance in VR Co-embodiment Environment

When evaluating our data and interview materials, we found that several participants preferred the condition where their own self-hands had ghost appearances while the co-embodied hands were realistic. They described the ghost hand's function as serving as guidance, which aligns with similar findings reported in previous work [9, 12, 14]. This feature appeared to influence users' attitudes towards co-embodiment tasks in VR, with the ghost hand appearance being perceived as an ancillary tool or an external intervention. It has been shown that participants referred to the ghost hand as an 'extension' that they possessed within the VEs, emphasizing its role as a supportive element rather than as part of their embodied self [12]. In addition, realistic hands, compared to hands with ghost appearances, are more attractive to users and provide more detailed feedback, as they are treated as distinct entities in the virtual space. These participants tended to choose realistic appearances for their self-hands and ghost appearances for the co-embodied hands, which they reported to be more effective for distinguishing roles. Some users from other research also reported realistic hands were more effective for completing tasks, as they felt these hands offered better precision and control compared to the intangible ghost hands [12].

Prior research on co-embodiment has shown that one user typically adopts the role of a learner, while another assumes that of a teacher [17, 31, 35, 53]. For learners, the shared avatar could reduce the availability of proprioceptive and motor feedback tied to their individual body movements, which is an essential component for effective motor skill acquisition [17, 35]. This reduction in self-hands movement could also negatively affect the learner's interaction with the co-embodied hand, as it limits their ability to identify and evaluate their own learning of the task [35, 53]. Thus, the introduction of another pair of self-hands, while initially appearing to add complexity, can actually provide additional sensory input through these self-hands' real movements. These extra signals could help re-calibrate the user's expectations and predictions about hand movements within the virtual space.

Although pairing two pairs of hands simultaneously introduces unique opportunities for collaborative interaction within the VE, it also presents a challenge. The existence of an extra virtual hand could lead to user confusion due to overlapping movements and visual similarities between these hands. This confusion could be problematic in motor learning contexts, where clear and consistent sensory input is critical to building accurate motor representations. When ghost hands are used for self-hands, users often perceive their role as an assistant rather than the primary agent of the task. This aligns with the experience of learners in co-embodiment scenarios, where the reduced connection between one's own movements and avatar's movements introduces ambiguity in task execution and role distinction [35, 56]. Conversely, realistic self-hands that provide more detailed feedback could better support motor learning by reinforcing users' SoO and SoA, thus enhancing their ability to understand and refine their movements [12]. Addressing these issues will require the careful design of visual or haptic feedback to ensure that both learners and teachers can effectively interpret their respective roles and contributions within a shared avatar.

6.3 Design Inspirations for Virtual Co-embodiment

The exploration of visual hand cues in VR co-embodiment offers rich insights for design considerations. Through the previous discussion, we summarize the following design inspirations.

Additional Self-hands Representation as the Visual Reference. The discrepancies between the movements of the co-embodied hand and the learner's actual hand can negatively impact the user's SoA and SoO in the shared avatar. This mismatch may hinder the learner's ability to accurately perceive their contributions using the shared avatar, potentially reducing the learning efficiency in motor skill acquisition tasks. To address this issue, additional self-hands could be introduced to reflect the learner's real hand movements directly. By providing this visual reference alongside the co-embodied hand, users are better able to distinguish their own movements.

Cognitive Load Reduction through Visual Factors. Extra visual cues using co-embodied avatars can provide valuable guidance, but should be carefully managed to avoid cognitive load. By providing more sensory feedback and improving the predictability of virtual interactions, users can engage more naturally and effectively with the virtual environment, potentially leading to improved task performance and greater user satisfaction. In our experiments, the hands with ghost appearance contributed to visual guidance, and their function was to decrease the predictive discrepancy during the task. Sharing one body with others is not usually a natural phenomenon in VR, and we have to be wary of introducing additional visual elements related to the human body, which may cause confusion.

Intention and Responsiveness in VR Co-embodiment Design. Designing for co-embodiment should prioritize synchronizing the self-hands with user's real movements, and also carefully managing the relationship between self-hands and the co-embodied hands. The potential to control and modify visual elements offers a powerful tool for designers to shape the virtual co-embodiment experience, such as entertainment, training, or therapeutic purposes. However, instead of striving for perfect visual representation, designers need to consider the characteristics of the task and whether the visual elements' setup will suit the context of the application. For example, in multi-person collaboration situations, social presence is one of the main factors that impact task experience.

7 Limitations and Future Work

This work had several limitations. First, the sample size was relatively modest, which may have limited the generalization of our findings. The second limitation is that we only tested the different visual representations with a 50% control weight. Varying the control weight higher or lower might reveal diverse effects when users face different visual combinations [35]. Furthermore, the cognitive confusion of self-hands and co-embodied hands occurs when two kinds of hands are too close, potentially leading to inaccurate judgments of the SoA and increasing cognitive load for the users. Besides, the study lacked objective performance metrics (e.g., task completion time or trajectory deviation) and did not include tasks of varying complexity, which could have provided deeper insights into how co-embodiment designs affect user experience under different conditions.

In future work, recruiting a larger participant pool can improve the robustness and generalization of the findings. Varying the control weights of co-embodied hands can alter user perceptions, potentially leading to differences in the SoA when self-hands and co-embodied hands are displayed simultaneously. Future studies can investigate the extent to which these visual representations influence the SoA in co-embodiment tasks with varying control weights [17, 33]. In addition, strategies to minimize cognitive confusion between self-hands and co-embodied hands, such as optimizing visual spacing or adding distinguishing visual cues, will be investigated to enhance user clarity and reduce cognitive load [4, 51, 55]. Moreover, incorporating objective performance metrics (e.g., task completion time, trajectory deviation) and designing tasks with varying complexities will be prioritized to provide more comprehensive insights into how co-embodiment designs perform. These insights are also valuable in various other applications using a co-embodied avatar, such as the learning of vocational skills [41] and rehabilitation [8, 11, 50, 63].

8 Conclusion

We investigated the effects of visual hand designs of co-embodied hands and self-hands within a virtual co-embodiment task on users' SoO, SoA, and social presence through two user studies. User Study One focused on how the existence of users' virtual self-hands on the co-embodied hands would influence these perceptions. It was found that users reported significantly higher SoO when self-hands were present. Therefore, self-hands were retained in User Study Two, which explored how *hand transparency* (ghost/realistic) and *hand congruency* (identical/different) would affect users' SoO, SoA, and social presence. Our results demonstrate that the identical hand congruency between the virtual self-hands and co-embodied hands, regardless of having different hand transparency levels, heightens SoO. In contrast, different hand transparency levels facilitated user differentiation, suggesting that more realistic hand appearances might increase their confidence in owning virtual hands, whether co-embodied hands or self-hands. Additionally, the transparency of virtual hands did not significantly impact SoA or social presence, likely due to the distinct conceptual definitions of each. These results suggest that visual factors of virtual hands (including the co-embodied hands and self-hands) have the potential to influence participants' subjective virtual interactive experience in co-embodiment tasks, and future studies could explore the implications of visual cues on user cooperative interactions in virtual co-embodiment settings.

9 Acknowledgment

We thank Duke Kunshan University for funding this research. The contributions of Jingjing Zhang, Hai-Ning Liang and Xin Tong were made while in their transition from Xi'an-Jiaotong Liverpool University and Duke Kunshan University to Hong Kong University of Science and Technology (Guangzhou), respectively.

References

[1] L. Afonso, P. Dias, C. Ferreira, and B. S. Santos. Effect of hand-avatar in a selection task using a tablet as input device in an immersive virtual environment. *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 247–248, 2017.

[2] F. Argelaguet, L. Hoyet, M. Trico, and A. Lécuyer. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *2016 IEEE Virtual Reality*, pp. 3–10. IEEE, 2016.

[3] S. Attasiriluk, A. Nakasone, W. Hantanong, R. Prada, P. Kanongchaiyos, and H. Prendinger. Co-presence, collaboration, and control in environmental studies: A second life-based approach. *Virtual Reality*, 13:195–204, 2009.

[4] M. Azmandian, M. S. Hancock, H. Benko, E. Ofek, and A. D. Wilson. Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016.

[5] A. P. Bartel and N. Sicherman. Technological change and the skill acquisition of young workers. *Journal of Labor Economics*, 16(4):718–755, 1998.

[6] L. Cardinali, C. Brozzoli, and A. Farne. Peripersonal space and body schema: two labels for the same concept? *Brain Topography*, 21(3):252–260, 2009.

[7] J. Casanueva and E. Blake. The effects of group collaboration on presence in a collaborative virtual environment. In *Virtual Environments 2000: Proceedings of the Eurographics Workshop in Amsterdam, The Netherlands, 2000*, pp. 85–94. Springer, 2000.

[8] Z. Chengjie and Y. Suiran. Virtual co-embodiment rehabilitation: An innovative method integrating virtual co-embodiment and action observation therapy in virtual reality rehabilitation. In *2024 17th International Convention on Rehabilitation Engineering and Assistive Technology*, pp. 1–6, 2024.

[9] W. Chinthammit, T. Merritt, S. Pedersen, A. D. Williams, D. C. Visentin, R. Rowe, and T. A. Furness. Ghostman: Augmented reality application for telerehabilitation and remote instruction of a novel motor skill. *BioMed Research International*, 2014.

[10] I. R. Cornford. Skill learning and the development of expertise. In *Adult Educational Psychology*, pp. 261–288. 2008.

[11] R. C. Curiel, T. Nakamura, H. Kuzuoka, T. Kanaya, C. Prahm, and K. Matsumoto. Virtual reality self co-embodiment: An alternative to mirror therapy for post-stroke upper limb rehabilitation. *IEEE Transactions on Visualization and Computer Graphics*, 30(5):2390–2399, 2024.

[12] D. Dewez, L. Hoyet, A. Lécuyer, and F. Argelaguet. Do you need another hand? investigating dual body representations during anisomorphic 3d manipulation. *IEEE Transactions on Visualization and Computer Graphics*, 28:2047–2057, 2022.

[13] C. Dubosc, G. Gorisse, O. Christmann, S. Fleury, K. Poinsot, and S. Richir. Impact of avatar facial anthropomorphism on body ownership, attractiveness and social presence in collaborative tasks in immersive virtual environments. *Computers & Graphics*, 101:82–92, 2021.

[14] M. Dürr, R. Weber, U. Pfeil, and H. Reiterer. Eguide: Investigating different visual appearances and guidance techniques for egocentric guidance visualizations. In *International Conference on Tangible, Embedded, and Embodied Interaction*, 2020.

[15] J. P. Freiwald, J. Schenke, N. Lehmann-Willenbrock, and F. Steinicke. Effects of avatar appearance and locomotion on co-presence in virtual reality collaborations. In *Proceedings of Mensch und Computer 2021*, pp. 393–401. 2021.

[16] R. Fribourg, F. Argelaguet, A. Lécuyer, and L. Hoyet. Avatar and sense of embodiment: Studying the relative preference between appearance, control and point of view. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):2062–2072, 2020.

[17] R. Fribourg, N. Ogawa, L. Hoyet, F. Argelaguet, T. Narumi, M. Hirose, and A. Lécuyer. Virtual co-embodiment: evaluation of the sense of agency while sharing the control of a virtual body among two individuals. *IEEE Transactions on Visualization and Computer Graphics*, 27(10):4023–4038, 2020.

[18] N. Getchell and J. Whitall. How do children coordinate simultaneous upper and lower extremity tasks? the development of dual motor task coordination. *Journal of Experimental Child Psychology*, 85(2):120–140, 2003.

[19] M. Gonzalez-Franco and T. C. Peck. Avatar embodiment. towards a standardized questionnaire. *Frontiers in Robotics and AI*, 5:74, 2018.

[20] T. Hagiwara, G. Ganesh, M. Sugimoto, M. Inami, and M. Kitazaki. Individuals prioritize the reach straightness and hand jerk of a shared avatar over their own. *Iscience*, 23(12), 2020.

[21] T. Hagiwara, M. Sugimoto, M. Inami, and M. Kitazaki. Shared body by action integration of two persons: Body ownership, sense of agency and task performance. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 954–955, 2019.

[22] C. Handford, K. Davids, S. Bennett, and C. Button. Skill acquisition in sport: Some applications of an evolving practice ecology. *Journal of Sports Sciences*, 15(6):621–640, 1997.

[23] H. Hapuarachchi, T. Hagiwara, G. Ganesh, and M. Kitazaki. Effect of connection induced upper body movements on embodiment towards a limb controlled by another during virtual co-embodiment. *Plos One*, 18(1):e0278022, 2023.

[24] H. Hapuarachchi and M. Kitazaki. Knowing the partner's objective increases embodiment towards a limb controlled by the partner. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 802–803, 2022.

[25] M. K. Harin Hapuarachchi. Knowing the intention behind limb movements of a partner increases embodiment towards the limb of joint avatar. *Scientific Reports*, 12, 2022.

[26] P. Heidicker, E. Langbehn, and F. Steinicke. Influence of avatar appearance on presence in social vr. In *2017 IEEE Symposium on 3D User Interfaces*, pp. 233–234, 2017.

[27] C. Heinrich, M. Cook, T. Langlotz, and H. Regenbrecht. My hands? importance of personalised virtual hands in a neurorehabilitation scenario. *Virtual Reality*, 25:313–330, 2021.

[28] A. Hiyama, Y. Doyama, M. Miyashita, E. Ebuchi, M. Seki, and M. Hirose. Artisanship training using wearable egocentric display. In *Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface Software and Technology*, pp. 81–82, 2011.

[29] D. Jo, K.-H. Kim, and G. J. Kim. Effects of avatar and background types on users' co-presence and trust for mixed reality-based teleconference systems. In *Proceedings the 30th Conference on Computer Animation and Social Agents*, pp. 27–36, 2017.

[30] E. Kal, R. Prosée, M. Winters, and J. Van Der Kamp. Does implicit motor learning lead to greater automatization of motor skills compared to explicit motor learning? a systematic review. *PloS One*, 13(9):e0203591, 2018.

[31] Y. Katsumata, Y. Inoue, S. Toriumi, H. Ishimoto, H. Hapuarachchi, and M. Kitazaki. Shared avatar for hand movement imitation: Subjective and behavioral analyses. *IEEE Access*, 11:96710–96717, 2023.

[32] K. Kilteni, R. Groten, and M. Slater. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, 2012.

[33] D. Kodama, T. Mizuho, Y. Hatada, T. Narumi, and M. Hirose. Enhancing the sense of agency by transitional weight control in virtual co-embodiment. In *2022 IEEE International Symposium on Mixed and Augmented Reality*, pp. 278–286. IEEE, 2022.

[34] D. Kodama, T. Mizuho, Y. Hatada, T. Narumi, and M. Hirose. Effect of weight adjustment in virtual co-embodiment during collaborative training. *Proceedings of the Augmented Humans International Conference 2023*, 2023.

[35] D. Kodama, T. Mizuho, Y. Hatada, T. Narumi, and M. Hirose. Effects of collaborative training using virtual co-embodiment on motor skill learning. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2304–2314, 2023.

[36] Z. Li, T. Huang, R. Patibanda, and F. Mueller. Ai in the shell: Towards an understanding of integrated embodiment. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–7, 2023.

[37] L. Lin and S. Jörg. Need a hand? how appearance affects the virtual hand illusion. In *Proceedings of the ACM Symposium on Applied Perception*, pp. 69–76, 2016.

[38] J.-L. Lugrin, M. Ertl, P. Krop, R. Klüpfel, S. Stierstorfer, B. Weisz, M. Rück, J. Schmitt, N. Schmidt, and M. E. Latoschik. Any "body" there? avatar visibility effects in a virtual reality game. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 17–24. IEEE, 2018.

[39] M. Luria, S. Reig, X. Z. Tan, A. Steinfeld, J. Forlizzi, and J. Zimmerman. Re-embodiment and co-embodiment: Exploration of social presence for robots and conversational agents. *Proceedings of the 2019 on Designing Interactive Systems Conference*, 2019.

[40] D. Mal, N. Döllinger, E. Wolf, S. Wenninger, M. Botsch, C. Wienrich, and

M. E. Latoschik. Am i the odd one? exploring (in) congruencies in the realism of avatars and virtual others in virtual reality. *Frontiers in Virtual Reality*, 5:1417066, 2024.

[41] F. Manuri, L. Sacco, F. De Pace, A. Sanna, and H. Kaufmann. A preliminary study of a co-embodiment approach for welding training. In *2024 IEEE 18th International Conference on Application of Information and Communication Technologies*, pp. 1–6. IEEE, 2024.

[42] M. Martini, K. Kilteni, A. Maselli, and M. V. Sanchez-Vives. The body fades away: investigating the effects of transparency of an embodied virtual body on pain threshold and body ownership. *Scientific Reports*, 5, 2015.

[43] M. Matamala-Gomez, T. Donegan, S. Bottiroli, G. Sandrini, M. V. Sanchez-Vives, and C. Tassorelli. Immersive virtual reality and virtual embodiment for pain relief. *Frontiers in Human Neuroscience*, 13:279, 2019.

[44] N. Ogawa, T. Narumi, and M. Hirose. Virtual hand realism affects object size perception in body-based scaling. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 519–528, 2019.

[45] P. N. Parmar, F. C. Huang, and J. L. Patton. Evidence of multiple coordinate representations during generalization of motor learning. *Experimental Brain Research*, 233:1–13, 2015.

[46] H. Pashler. Dual-task interference in simple tasks: data and theory. *Psychological Bulletin*, 116(2):220, 1994.

[47] T. C. Peck and M. Gonzalez-Franco. Avatar embodiment. a standardized questionnaire. *Frontiers in Virtual Reality*, 1:575943, 2021.

[48] A. M. Persky and J. D. Robinson. Moving from novice to expertise and its implications for instruction. *American Journal of Pharmaceutical Education*, 81(9):6065, 2017.

[49] J. Pinkl and M. Cohen. Vr drumming pedagogy: Action observation, virtual co-embodiment, and development of drumming "halvatar". *Electronics*, 12(17):3708, 2023.

[50] I. Podkosova and H. Brument. Towards full body co-embodiment of human and non-human avatars in virtual reality. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 432–435. IEEE, 2024.

[51] A. Pusch, O. Martin, and S. Coquillart. Hemp-hand-displacement-based pseudo-haptics: A study of a force field application. *2008 IEEE Symposium on 3D User Interfaces*, pp. 59–66, 2008.

[52] M. V. Sanchez-Vives, B. Spanlang, A. Frisoli, M. Bergamasco, and M. Slater. Virtual hand illusion induced by visuomotor correlations. *PloS One*, 5(4):e10381, 2010.

[53] H. Takita, Y. Hatada, T. Narumi, and M. Hirose. Effects of virtual co-embodiment on declarative memory-based motor skill learning. In *ACM Symposium on Applied Perception 2023*, pp. 1–10, 2023.

[54] H. Takita, D. Kodama, Y. Hatada, T. Narumi, and M. Hirose. Motor skill learning by virtual co-embodiment with an ai teacher trained in human teaching behavior. In *ACM Symposium on Applied Perception 2024*, pp. 1–9, 2024.

[55] Y. Tao, C. Y. Wang, A. D. Wilson, E. Ofek, and M. González-Franco. Embodying physics-aware avatars in virtual reality. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023.

[56] Y. Tian, Y. Bai, S. Zhao, C.-W. Fu, T. Yang, and P.-A. Heng. Virtually-extended proprioception: Providing spatial reference in vr through an appended virtual limb. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

[57] M. Tombu and P. Jolicœur. A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, 29(1):3, 2003.

[58] K. P. Venkatraj, W. Meijer, M. Perusquia-Hernandez, G. Huisman, and A. El Ali. Shareyourreality: Investigating haptic feedback and agency in virtual avatar co-embodiment. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2024.

[59] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 143–146, 2011.

[60] U. Yang and G. J. Kim. Implementation and evaluation of just follow me: An immersive, vr-based, motion-training system. *Presence: Teleoperators & Virtual Environments*, 11:304–323, 2002.

[61] B. Yoon, H.-i. Kim, G. A. Lee, M. Billinghurst, and W. Woo. The effect of avatar appearance on social presence in an augmented reality remote collaboration. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 547–556, 2019.

[62] B. Yoon, J.-e. Shin, H.-i. Kim, S. Young Oh, D. Kim, and W. Woo. Effects of avatar transparency on social presence in task-centric mixed reality remote collaboration. *IEEE Transactions on Visualization and Computer Graphics*, 29(11):4578–4588, 2023.

[63] J. Zhang, M. Huang, Y. Chen, K.-L. Liao, J. Shi, H.-N. Liang, and R. Yang. Touchmark: Partial tactile feedback design for upper limb rehabilitation in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 30(11):7430–7440, 2024.

[64] J. Zhang, M. Huang, R. Yang, Y. Wang, X. Tang, J. Han, and H.-N. Liang. Understanding the effects of hand design on embodiment in virtual reality. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 37:e10, 2023.