

Stretch your reach: Studying Self-Avatar and Controller Misalignment in Virtual Reality Interaction

Jose Luis Ponton

jose.luis.ponton@upc.edu

Universitat Politècnica de Catalunya
Barcelona, Spain

Alejandro Beacco

alejandro.beacco@upc.edu

Universitat Politècnica de Catalunya
Barcelona, Spain

Reza Keshavarz

reza.keshavarz@studio.unibo.it

Università di Bologna
Bologna, Italy

Nuria Pelechano

npelechano@cs.upc.edu

Universitat Politècnica de Catalunya
Barcelona, Spain

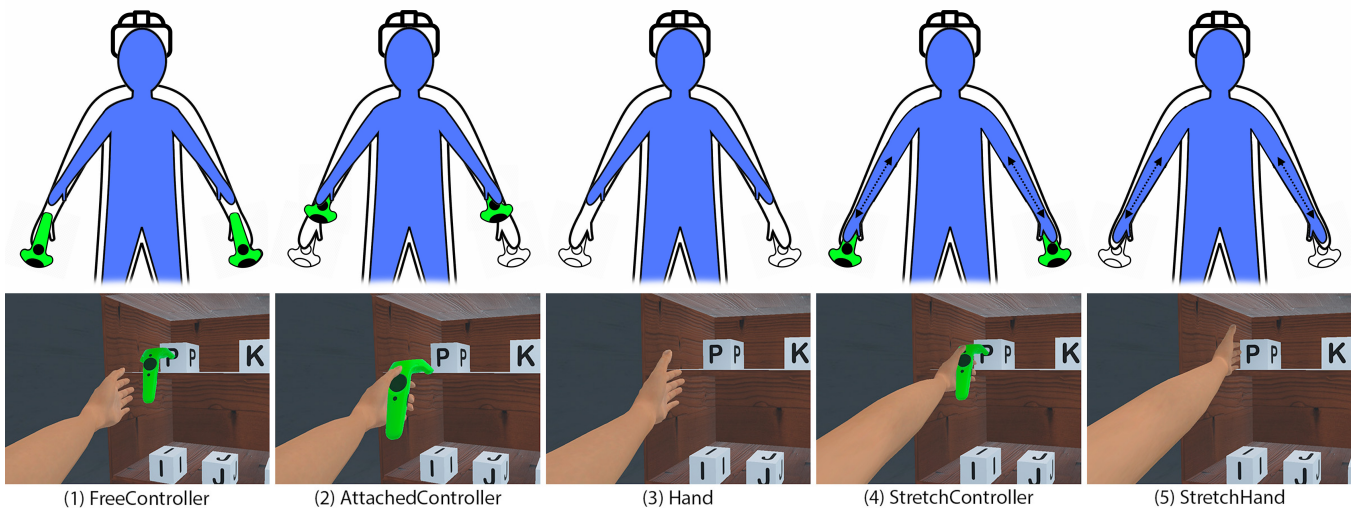


Figure 1: Five interaction modes are proposed to manipulate virtual objects when the user has a full-body avatar animated with Inverse Kinematics. (1) *FreeController*, the controller is rendered and may appear to fly when the avatar's arm is shorter than the user's arm; (2) *AttachedController*, the controller is rendered always attached to the user's virtual hand even when the avatar's arm is shorter than the user's arm; (3) *Hand*, the controller is not rendered and thus the virtual hand reach depends on the avatar's arm length; (4) *StretchController*, the controller is rendered and the arms stretch to always match the position of the virtual and real hand, thus always holding the virtual controller; similarly (5) *StretchHand* has the same behavior but without rendering the controller. For a summary of all methods see Table 1.

ABSTRACT

Immersive Virtual Reality typically requires a head-mounted display (HMD) to visualize the environment and hand-held controllers to interact with the virtual objects. Recently, many applications display full-body avatars to represent the user and animate the arms to follow the controllers. Embodiment is higher when the self-avatar movements align correctly with the user. However, having a full-body self-avatar following the user's movements can be challenging due to the disparities between the virtual body and the user's body. This can lead to misalignments in the hand position that can be noticeable when interacting with virtual objects. In this work, we propose five different interaction modes to allow the user to interact with virtual objects despite the self-avatar and controller misalignment and study their influence on embodiment,

proprioception, preference, and task performance. We modify aspects such as whether the virtual controllers are rendered, whether controllers are rendered in their real physical location or attached to the user's hand, and whether stretching the avatar arms to always reach the real controllers. We evaluate the interaction modes both quantitatively (performance metrics) and qualitatively (embodiment, proprioception, and user preference questionnaires). Our results show that the stretching arms solution, which provides body continuity and guarantees that the virtual hands or controllers are in the correct location, offers the best results in embodiment, user preference, proprioception, and performance. Also, rendering the controller does not have an effect on either embodiment or user preference.

CCS CONCEPTS

• **Human-centered computing** → **User studies; User models; Computing methodologies** → **Virtual reality; Perception.**

KEYWORDS

virtual reality, 3D interaction, avatars, perception, embodiment

1 INTRODUCTION

Immersive Virtual Reality is a technology with the potential to revolutionize fields like education, healthcare, and entertainment, offering immersive, impactful experiences. Having users embodied with a virtual avatar can improve spatial perception, among others, and has already been used for applications in learning, rehabilitation, and mental health. Immersive VR can make the users feel as if they are actually present in a virtual world and behave as if they were in the real world. A virtual body representation that follows the user's moves with visuomotor synchrony can provide the illusion of body ownership, leading to embodiment [18], which can enhance presence [38]. Ideally, the virtual avatar should perfectly mimic the user to facilitate natural interactions since inaccuracies in body pose can reduce embodiment and performance [45]

Given the challenges of having accurate avatar representations when wearing a Head Mounted Display (HMD), often, the user is represented with a floating upper torso and hands. In these cases, the upper body is located under the HMD position, and the virtual hands are rendered at the position of the handheld controllers, and provide a way to interact with the virtual world. However, when rendering a full virtual body, it is often difficult to accurately co-locate the virtual and real hand positions due to limitations of the virtual body, such as body size or proportions [33], skeleton simplification, or inaccurate offsets with respect to trackers [32]. Respecting the virtual body dimensions may lead to misalignments between the virtual hand/controller and the real hand/controller when reaching far-away objects.

Body discrepancies may not be perceptually noticeable in applications that do not require careful manipulation of objects, for example, those where the user has to move freely [32], walk [23], gesticulate while talking [39], or even hold a controller and eventually press a button [27]. However, when direct manipulation is needed while rendering a full avatar, it may be necessary to redirect the hand position to respect the limits imposed by the virtual body. Virtual hand redirection, when rendering only hands or hands with part of the arm, has been studied to improve ergonomics [26], to induce the perception of weight in virtual objects [36], or to induce a proprioceptive drift and a change in body schema [20]. However, when a full virtual body is being rendered, altering the position of the virtual body with respect to the user could affect embodiment. The work by Tao et al. [42] improved the realism of the interactions between the self-avatar and the environment by adding physics to the virtual body. This led to discrepancies between the virtual and the real body pose, which could increase or reduce embodiment depending on the intensity. This work set the basis to continue researching to what extent we can introduce discrepancies to improve interaction in the virtual world without reducing embodiment. In this paper, we focus on interactions between the hand and the virtual objects.

Current solutions that can be found in most applications either render only floating hands always on the controller position [1–3], or not render the controllers when having a full avatar [7, 11, 21], to avoid the misalignments with the virtual hands caused by the limitations imposed by having a full-body avatar.

Not rendering the controllers may be enough for VR applications with little manipulation of objects, such as applications to imitate an avatar to learn to dance or practice yoga, or social applications providing non-verbal communication. However, if the virtual controllers are required for interaction purposes (e.g., used as a tool to interact closely with objects), or, similarly, if a virtual tool must be rendered, then certain trade-offs arise. For example, suppose the controllers are rendered in their real-world position, and the user has a full-body avatar. In that case, users may be trying to interact with a virtual object that is far away (users need to stretch the arm) and observe the virtual controller moving away from the avatar's hand [33] (as if it was flying because the virtual arm cannot reach such position). This can be perceptually noticeable for users as they may notice an inconsistency between haptic feedback (feeling the controller's touch in the hand) and visual feedback (observing the controller flying away from the virtual hand).

One solution could be to render the controller always attached to the avatar's hand to achieve multisensory feedback (visual and haptic). Then, by the induced embodiment, we may convince the users that their body has changed dimensions to accept the modified body as their own [20]. It could then be possible to have believable interaction with the virtual objects in the environment despite the misalignments between the virtual body and the user. In some studies, it has been observed that this may lead to a proprioceptive drift which makes the user move toward the avatar pose to fill the gap [13]. However, when trying to grab out-of-reach objects, it is impossible to close the gap and may lead to a proprioceptive conflict. Another alternative could be to stretch the avatar's arm so that the virtual hand can always hold the controller [8], although the unnatural skinning could negatively affect appearance, which is also important for embodiment [1, 9].

Our goal is to study interaction metaphors that can be used as solutions to overcome current limitations when animating a full virtual avatar to represent the user while interacting with the environment. In this work, we propose and explore five interaction metaphors consisting of different representations for the user when interacting with virtual objects. These representations are different combinations of binary factors, such as rendering the virtual controllers, attaching them to the virtual hands, and stretching the virtual arms. Our specific research questions are the following:

- **RQ1:** How does the rendering of virtual controllers impact user embodiment and performance in object manipulation tasks within virtual reality?
- **RQ2:** How does the accuracy of virtual controllers' absolute location affect user embodiment and performance in object manipulation tasks within virtual reality?
- **RQ3:** How does attaching the virtual controllers to the virtual hand impact user embodiment and performance in object manipulation tasks within virtual reality?

The goal of our study is to determine which of our interaction metaphors is preferred by users and can enhance embodiment and

performance. All our methods have the same avatar dimensions corresponding to a uniform scaling of an avatar based on the participant’s height, and, thus, is not an exact match of the participant’s proportions. The contributions of this work are:

- A study of five distinct interaction metaphors, focusing on three key factors: *Controller*, *Attached*, and *Stretch*. We analyze their relationship with embodiment, performance, user preference, and proprioception in virtual reality. A significant finding from our study is that stretching arms consistently yields the best results across all these metrics. Furthermore, visually rendering VR controller does not significantly impact these outcomes.
- Based on the insights from the study, we formulate a set of guidelines for designing manipulation techniques for full-body self-avatars in virtual reality.

2 RELATED WORK

Head-mounted displays for Virtual Reality are increasingly becoming popular for performing some tasks due to the exceptional degrees of immersion they afford. Specifically, tasks involving high interactivity levels seem to improve the sense of presence, thereby enhancing the overall user experience [44]. There is extensive literature focused on discerning methods to quantify embodiment, presence, and immersion in VR. Probably the most relevant of those studies is the rubber hand illusion [2], which examines the sense of ownership over a rubber hand synchronized with visual and tactile stimuli, leading to the illusion that the rubber hand belongs to the participant. This feeling of ownership can also be replicated in a virtual environment [40], even without tactile stimulation [37].

Seeking to better understand and classify the mechanisms that generate these illusions, Kilteni et al. [18] delves into the Sense of Embodiment (SoE) concept, defined as perceiving an external body as one’s own. They propose three key dimensions influencing SoE: the Sense of Self-Location (SoL) pertaining to the feeling of being within the virtual body, the Sense of Agency (SoA) regarding the perception of control over the virtual body, and the Sense of Body Ownership (SoO) about the feeling that the virtual body is one’s own.

Given the growing interest in the SoE, Gonzalez-Franco and Peck [14] conducted a comprehensive review of embodiment questionnaires, resulting in a standardized questionnaire for embodiment experiments. This questionnaire encompasses six primary question types to assess embodiment, namely, body ownership, agency and motor control, tactile sensations, body location, external appearance, and response to external stimuli. Later, Peck and Gonzalez-Franco [31] further refined this questionnaire, proposing four key embodiment components: appearance, response, ownership, and multi-sensory. While agency, touch, and localization are not explicitly included in this categorization, the researchers discovered that these components contribute to the four major embodiment categories as they correlate with other senses.

2.1 Body Representation in Virtual Reality

While numerous factors, such as the visual rendering of the virtual environment or the degree of interactivity [44], influence the SoE, a crucial contributor is the representation of the user’s body via

an avatar—commonly termed a self-avatar. This representation considerably augments SoE and proves advantageous for different tasks like egocentric distance estimation, spatial reasoning, and collision avoidance [29, 30, 35].

A variety of studies have examined how different aspects of a full-body avatar impact SoE. For instance, Gonçalves et al. [11] analyzes the influence of varying numbers of tracking points and locomotion, whereas Fribourg et al. [9] investigates the effect of point of view, appearance, and control on SoE and its components. However, the impact of the interaction between the full-body avatar and the virtual objects on SoE remains unexplored.

To address this gap, several libraries have been introduced to augment embodiment with full-body avatars [28, 33]. In our study, we go one step further: we instantiate the user in a full-body self-avatar to create a convincing embodiment illusion and then examine how the representation of virtual controllers, hands, and their positioning in the virtual environment—in relation to their real counterparts—influence interaction and immersion.

2.2 Redirected hand position

Hand redirection has been used in immersive VR with several goals. Rietzler et al. [36] introduced perceivable tracking offsets between the user’s hand and the virtual hand in order to simulate weight when carrying objects. Montano Murillo et al. [26] developed a redirection technique to improve ergonomics by reducing the range of physical movements needed to reach virtual objects within a virtual environment. Their method takes advantage of the dominance of the human visual system over the proprioceptive system. Hoyet et al. [16] presented two methods to recover the position mismatch between users’ real and virtual hands after releasing contact with a virtual object. Gonzalez and Follmer [12] proposed a sensorimotor model of hand redirection to simulate several movement features such as hand trajectory or velocity. Their model could be used as a tool to evaluate hand redirection techniques without the need for user testing.

The main difference between previous work on redirected hand position is that they render either just a hand or a hand with part of the arm. Still, without a full virtual body or else the virtual body is mostly static and only the hand can perform small movements. Therefore, they do not need to worry about embodiment or body continuity when the user has complete freedom of movement. The challenge when rendering a hand in a different position while having a full avatar is that it may break the body ownership illusion, which will affect embodiment. Previous work has shown that when elongating the arm length, users can feel body ownership, although the illusion diminishes with the length of the arm [20]. The effect is induced by the congruence visuo-motor and visuo-tactile feedback, even if there are incongruences between proprioception and visual feedback (i.e., the hand being rendered far away from the real hand). However, in their work, the user could only sit still resting the arm on a table while doing small movements to touch the surface.

2.3 Interaction Modes in Virtual Reality

Interaction in VR is crucial to enhance the levels of immersion, necessitating meticulous design to bolster user performance within virtual environments. Typically, VR interaction is facilitated by

monitoring the user’s hand movements via a device, such as a VR controller or a vision-based hand-tracking mechanism.

Early research conducted by Argelaguet et al. [1] explored the impact of virtual hand rendering on the SoE for a pick-and-place task. The results indicated that a higher SoA was achieved with more simplistic models, such as a sphere, as compared to a realistic hand model. However, the potential inaccuracy in hand tracking may have influenced these findings. In line with this, several subsequent studies compared multiple methods to render a virtual hand. Lin and Jörg [22] found that all hand rendering types could induce an illusion of body ownership, although less effectively, with non-anthropomorphic models. Ricca et al. [34] examined the role of hand visualization in tool-based training, finding no significant difference in performance when hands were rendered versus when they were not.

Subsequently, some researchers shifted their focus to the rendering of the virtual arm. Tran et al. [43] reported that more basic representations (such as only hand or hand and wrist) performed faster in their tasks, yet no significant differences were found in accuracy, SoA, and SoO. Kober et al. [21] explored the EEG activity when rendering a realistic hand and arm representation versus a skeleton-based one, finding that similar activations were achieved with the more realistic arm model.

Despite these advancements, research into hand or arm rendering has led to varied results, suggesting the type of task being performed may be a critical factor, and thus, more research in this area is needed. Given the potential influence of hand tracking accuracy on these outcomes [7], some studies have compared hand tracking systems with 6-DoF virtual reality controllers [4–6, 15]. Here, all studies found that the use of VR controllers enhanced performance, user experience, and embodiment.

Although VR controllers are commonly used for VR interaction, there are few studies on their effects on the participants. There is, therefore, a need to expand knowledge about VR interactions involving controllers. In response, Gao and Boehm-Davis [10] developed a general questionnaire for evaluating interactions with objects across various VR applications. Lougiakis et al. [24] compared three methods of rendering controllers—a sphere, a virtual controller, and a hand. They found no significant differences in SoA; however, the sphere under-performed significantly, and the controller outperformed the other conditions in the positioning task. Other studies compared the use of VR controllers for rehabilitation [17], although their focus was predominantly on the physical controller rather than the virtual representation.

In our work, we focus on studying interaction metaphors when using controllers with a full-body avatar. Our interaction metaphors propose alternatives to solve the mismatches between the user and the avatar, due to the simplified avatar skeleton with respect to the human counterpart.

3 DESIGN

In this section, we present the different factors involved in interacting with virtual objects that we have considered to examine. By manipulating these factors, we will obtain the different interaction modes that will be part of the user study.

3.1 Interaction modes

When embodying the user into a full-body avatar while using hand-held controllers, there are different decisions to be made in order to achieve different embodiment configurations. Depending on the factors we consider and those decisions, some problems might arise or not. For instance, when interacting with virtual objects, we often consider whether we should render the virtual representation of the VR controllers. While rendering the VR controllers provides accurate visual feedback, it may be detrimental if it brings up problems with virtual arm length or hand position due to size differences between the user and the avatar. Also, if the virtual controller consistently stays rendered in the palm of the virtual hand, we are able to maintain multi-sensory feedback and body continuity. However, this may be at the expense of not correctly aligning the virtual and the real controllers when the avatar’s arms are not long enough, which can lead to confusion due to conflicting proprioceptive feedback (the user can sense that the distance to the hand is not the real one). An option to solve this problem when the real controller is located outside the avatar’s virtual arm reach is to stretch the arm. This provides virtual body continuity and allows the virtual controllers to be correctly aligned with the real ones.

Therefore, we have considered and examined in our study the following three critical factors:

- (a) *Controller*: whether the virtual controller is visually represented or not.
- (b) *Attached*: whether the virtual controller remains consistently positioned in the palm of the virtual hand.
- (c) *Stretch*: whether the avatar’s arm is extended to align the virtual controller with its corresponding absolute position in the physical world.

These three binary factors (see Fig. 2) are used to guide the design of our interaction metaphors that we have devised and implemented to study the research questions outlined in Section 1. The manipulation of the *Controller* factor directly influences the study of **RQ1**. However, the remaining factors contribute to both **RQ2** and **RQ3**. The *Attached* factor directly impacts **RQ2**; nonetheless, forcing the virtual controller always to be attached to the palm of the virtual hand may introduce misalignment issues as the avatar’s arm may not perfectly match the user’s arm. To address this concern, we introduced the *Stretch* factor, which allows us to maintain the accurate absolute positioning of the virtual controller while also ensuring its attachment to the hand.

Considering all possible combinations of these three factors would result in eight distinct interaction modes. However, we carefully selected five interaction modes to ensure validity within the VR domain. Hence, there are three invalid combinations: firstly, not attaching the controller to the hand when the arm is stretched; secondly, when the controller is not rendered it does not matter whether the virtual controller is attached or not, consequently eliminating two more modes.

The final five interaction modes (*Mode*) are summarized in Table 1 and visualized in Fig. 1. (1) *FreeController* renders the controller in its real-world position but can break body continuity and visual-tactile sensory feedback. (2) *AttachedController* positions the controller in the avatar’s hand, ensuring body continuity and multi-sensory cues, but may not be in its real-world position thus

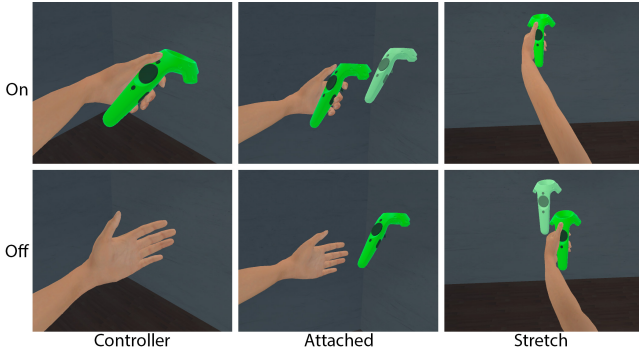


Figure 2: Illustration of the different factors in two scenarios. The top row represents the scenario in which the factor is enabled, and the bottom row corresponds to the disabled factor. From left to right, the factors *Controller*, *Attached*, and *Stretch* are shown. In situations where the virtual controller is not aligned with the real one, the real controller is depicted in the figure with a lighter color (not in the simulation).

leading to a proprioceptive conflict. (3) *Hand* mode, where the VR controller is not rendered, might not always align with the real-world hand position due to the virtual avatar’s arm being shorter than the user’s, thus also leading to a proprioceptive conflict. (4) *StretchController*, stretches the avatar’s arm to reach the real-world controller location when necessary and always renders the controller attached to the hand. Finally, (5) *StretchHand*, is similar to *StretchController*, but without rendering controller.

Although not initially employed to design the interaction modes, we introduce a fourth auxiliary binary factor to facilitate the study of proprioception for our interaction modes:

- (d) *Location* determines when the virtual controller is rendered, whether its position remains aligned with its corresponding absolute position in the physical world. If the virtual controller is not rendered, *Location* refers to whether the virtual hand position was aligned or not with the real hand position.

The inclusion of the *Location* factor will serve to isolate the impact of the precise virtual controller or hand positioning, independent of other factors that primarily affect rendering aspects.

Mode	Factor			
	<i>Controller</i> (a)	<i>Attached</i> (b)	<i>Stretch</i> (c)	<i>Location</i> (d)
(1) <i>FreeController</i>	Yes	No	No	Yes
(2) <i>AttachedController</i>	Yes	Yes	No	No
(3) <i>Hand</i>	No	N/A	No	No
(4) <i>StretchController</i>	Yes	Yes	Yes	Yes
(5) <i>StretchHand</i>	No	N/A	Yes	Yes

Table 1: Interaction modes used in the study. Four binary factors are used to determine the behavior of each mode: (a) *Controller*, (b) *Attached*, (c) *Stretch*, and (d) *Location*. The interaction modes can be seen in Fig. 1.

3.2 Interaction Tasks

We designed three distinct interaction tasks encompassing various upper-body actions to evaluate task performance and user immersion in the self-avatar. These tasks aimed to provide users with various upper-body interactions manipulating virtual objects and enable them to assess the level of immersion achieved.

Cube Task. Participants pick up five cubes with different alphabet letters printed on them and place them on corresponding docks. The cubes are initially positioned on a shelf within the virtual room. Participants must use their virtual hand or controller (depending on the mode) to grasp the cubes by pressing the trigger button and moving them to the correct dock that matches the letter printed on the cube. The cubes are deliberately placed at varying heights on the shelf to simulate different difficulty levels in picking them up, encouraging participants to extend their arms when necessary. This task assesses various VR interactions involving object manipulation and spatial movement within the virtual scene.

Cannon Task. Virtual balls are launched into the virtual environment by different cannon shooters which users must catch. Participants are positioned behind a virtual line, and the cannon shooters appear on a wall before them. A total of twenty-four balls are sequentially shot, each with a slightly different direction and velocity. Participants must attempt to catch the balls using their virtual hands or controllers (depending on the mode), without the need for pressing any buttons—simply by touching the balls. A sound accompanies successful catches, while missed catches trigger a distinct, identifiable sound. This task assesses fast-paced mid-air interactions, requiring quick reflexes and hand-eye coordination.

Painting Task. Participants had to draw simple shapes on a virtual whiteboard. A template of basic figures is displayed on the whiteboard, including a triangle, a square, and a circle. Using a virtual brush or the tip of their virtual index finger (in the absence of controllers), participants approach the whiteboard surface to create strokes and trace the outlines of the given shapes. This task assesses participants’ precision in interacting with the virtual environment by following the template and accurately reproducing the shapes through drawing actions.

4 USER STUDY

To address our research questions, we conducted a within-subjects study encompassing five conditions corresponding to the five different interaction modes in VR. While the primary independent variable in our study is the interaction modes, we also analyze the three factors—*Controller*, *Attached*, and *Stretch*—as separate independent variables to discover additional insights.

4.1 Apparatus

The experiments took place within a laboratory room measuring 6 m x 6 m in size. The average duration of the experiment was approximately 45 minutes. The virtual environment was developed using Unity 2021.3 and executed on a PC equipped with an Intel Core i7-12700k CPU, an Nvidia GeForce RTX 3090 GPU, and 32 GB of RAM. For the VR experience, an HTC Vive Pro HMD with a resolution of 1440 x 1600 pixels per eye, a field of view of 110°,

and a refresh rate of 90 Hz was utilized. An external battery and a wireless adapter were employed for the HTC Vive system to enhance freedom of movement, eliminating the need for cables. The participants’ pelvis and feet were tracked using three 6-DoF HTC Vive trackers 3.0, while two HTC Vive controllers were held in the participants’ hands. In order to minimize occlusions, four SteamVR Base Station 2.0 units were installed at each corner of the room.

4.2 Participants

A total of 40 participants took part in the study: 34 were right-handed and 6 left-handed, with a gender distribution of 19 females and 21 males. Most participants were university students between the ages of 18 and 24. Participants were not compensated. Regarding gaming experience, 14 participants reported high, 13 medium, 10 low, and 4 no experience. Regarding VR experience, 6 reported high, 3 medium, 18 low, and 14 no experience. Nearly all participants were familiar with using computers, although only a few had prior exposure to VR technology.

The Ethics Committee of the Universitat Politècnica de Catalunya (UPC) issued a favorable opinion on the ethical aspects related to the research carried out in this project (ID Code: 2023.04). The favorable approval of the application implies that the reviewed project complies with the criteria established by the institution’s own expertise.

4.3 Design

An overview of the user study protocol can be seen in Fig. 3. At the beginning of the study, participants were embodied in a virtual avatar animated using the *AvatarGo* library [33], which was designed to induce avatar embodiment. Having a well-calibrated avatar already produces high levels of embodiment and is beneficial for interaction in VR [7]. *AvatarGo* employs Unity’s built-in IK solver to animate the avatar’s limbs (2-segment kinematic chains). However, their method does not solve each limb considering the full-body pose. Instead, it computes the positions of each limb’s joints independently, targeting one end-effector per limb. To enhance the overall body pose, *AvatarGo* integrates forward kinematics (FK) for two additional joints—the head and spine. This allows the self-avatar to perform more complex movements, such as leaning forwards and sideways, contributing to a more realistic and dynamic representation. Therefore, participants could see and move through the virtual world as if they were controlling a virtual person with their own movements, and focus on the interaction techniques rather than other embodiment issues. Furthermore, participants were given sufficient time to adjust and become accustomed to their new virtual avatars by following a series of standard embodiment instructions.

Throughout the first part of the experiment, for each condition, participants played in the virtual environment the three different tasks explained in Section 3.2: the *Cube*, *Cannon* and *Painting* tasks. The order of the conditions was established by a 10 x 5 balanced Latin square, and within each condition, the order in which the tasks were performed was randomized. After each condition, an embodiment questionnaire (see Table 2) was displayed inside the virtual environment as shown in Fig. 3. Questions were based on the

EQ1	I felt as if the movements of the virtual body were influencing my own movements.
EQ2	I felt like I could control the virtual body as if it were my own body.
EQ3	It seemed as if the touch I felt was caused by the controllers/block/brush touching the virtual body.
EQ4	It seemed as if my body was touching the controller/block/brush.
EQ5	I felt as if my body was located where I saw the virtual body.
EQ6	It seemed as if I felt the touch of the controller/block/brush in the location where I saw the virtual body touched.

Table 2: Embodiment Questionnaire. Participants had to score from 1 to 7 on the statements where 1 means they strongly disagree and 7 means they agree completely.

PQ1	How much you liked the interaction technique? (Rate each of the interaction modes from 0 to 100)
PQ2	How easy was it to manipulate the objects with this technique? (Rate each of the interaction modes from 0 to 100)

Table 3: Preference Questionnaire. Participants had to score from 0 to 100 for each interaction mode for two questions .

work by Peck and Gonzalez-Franco [31] in which they propose four key embodiment components: appearance, response, ownership, and multi-sensory. We selected a subset of six questions to suit the specific needs of our experiment, with the aim of avoiding participant fatigue considering the duration of the experiment. While the chosen questions cover all the proposed four components of the SoE, they were particularly focused on those highly influencing ownership and multi-sensory aspects.

Following the three tasks across all conditions, we conducted an additional iteration of the conditions. During this phase, participants were queried about their perception of any differences between the positions of their real hands and those of the virtual hands or controllers. To gather this information, a proprioceptive questionnaire was presented within the virtual environment, allowing participants to provide binary responses for each condition to the following question **LQ1**: *Do you think your hand was in the position of your virtual hand?*

At this stage, the experimenter explained the various interaction techniques to the participants. Following this, participants had the opportunity to freely explore and experience all the conditions while providing their feedback through a final preference questionnaire inside the VR (see Table 3 and Fig. 4 right). Notice that participants could change between different modes as many times as they wished using the VR controller’s touchpad and a virtual watch (see Fig. 4 left), thus, they could easily compare between interaction modes and rate them based on preference. This questionnaire aimed to capture participants’ personal preferences regarding the different interaction techniques used in the study.

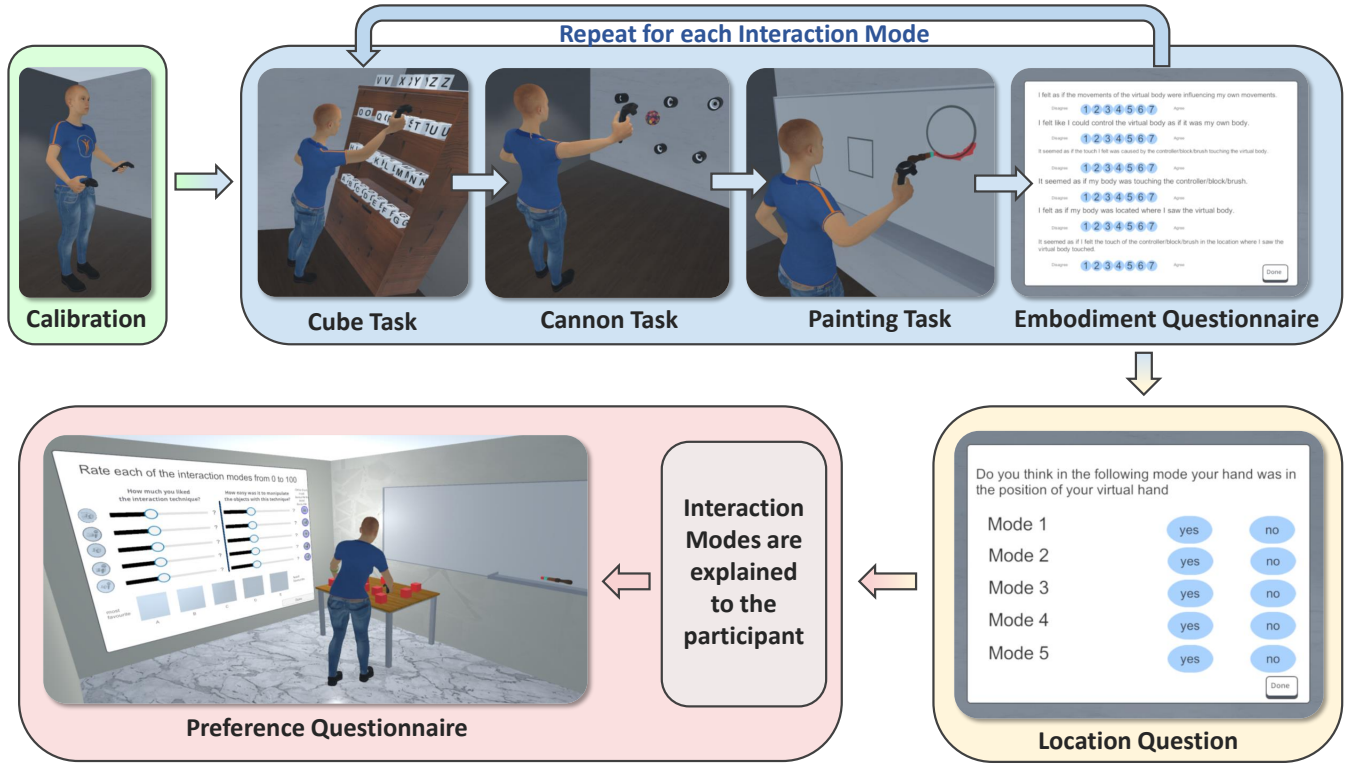


Figure 3: Diagram of the user study protocol. The experiment begins with participants being embodied in a virtual avatar, animated using the *AvatarGo* library [33]. Participants then perform three tasks (in randomized order) in the virtual environment and answer an embodiment questionnaire for each condition. In the subsequent stage, participants answer a question regarding their perception of the positions of their real hands and virtual hands. Finally, after being introduced to the interaction modes, participants explore all conditions freely and provide feedback via the final preference questionnaire.

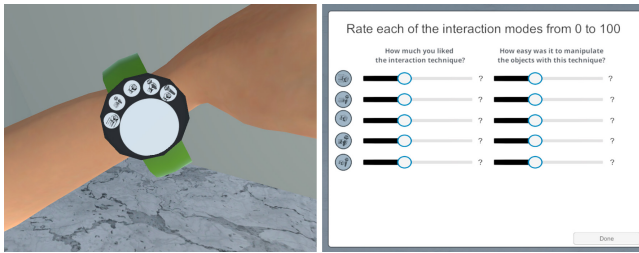


Figure 4: In the preference questionnaire phase, participants could freely switch between interaction methods using the VR controller’s touchpad interfaced with a virtual watch display (left), aiding users in visualizing their selected mode. To simplify the recognition of each interaction mode, they were depicted by a unique icon, consistent with its representation in the preference questionnaire (right).

4.4 Measures

In this section, we detail the variables used in the subsequent analysis of the results. First, as suggested by Peck and Gonzalez-Franco [31], we aggregate the embodiment questions to derive the final embodiment score $Embodiment = EQ1 + EQ2 + EQ3 + EQ4 + EQ5 + EQ6$

(see Table 2). Similarly, we calculate user preference as $Preference = PQ1 + PQ2$ (see Table 3). Lastly, we define *Proprioception* to represent whether users perceived the virtual hand to align with the real one. This binary variable is directly derived from question LQ1 (see Section 4.3).

Following this, we delineate the performance metrics used in each task:

- $Performance_{Cube}$ corresponds to the completion time from the start of the *Cube Task* until the final cube is positioned on the table.
- $Performance_{Cannon}$ represents the count of balls correctly caught during the *Cannon Task*.
- $Performance_{Painting}$ measures the similarity between the overlap user-drawn texture D with the ground truth G in the *Painting Task*. It is computed as the intersection over union of drawn pixels, expressed as $Performance_{Painting} = \frac{G \cap D}{G \cup D}$.

Finally, we aggregate the rank transformations of each task’s performance metric to compute an overall performance score:

$$Performance = \frac{1}{3} (\text{rank}(Performance_{Cube}) + \text{rank}(Performance_{Cannon}) + \text{rank}(Performance_{Painting})). \quad (1)$$

Here, for each user, the rank transformation converts numerical values into ranks for each condition, assigning values from 1 (for the least performant condition) to 5 (for the most performant condition).

4.5 Hypotheses

In this section, we aim to provide a comprehensive presentation of our hypotheses, outlining the specific aspects of the user study that will be analyzed. Our *a priori* hypotheses are based on previous related work, and are systematically categorized based on the factors under investigation, thereby facilitating the understanding of the results.

4.5.1 Embodiment. Our first set of hypotheses is based on early work on the effect of visual-sensorimotor contingencies, visual features, and proprioception on the SoE [2, 19, 20, 24, 25, 37]. **H1** suggests that aligning visual, tactile, and proprioceptive stimuli enhances the SoE. Therefore, rendering VR controllers and matching their position (absolute and relative) increases embodiment. Because of that, we hypothesize that not all modes will convey the same level of embodiment (**H2**). Finally, we also speculate that we can positively affect embodiment if the user believes the controller position is correct (**H3**).

H1 *Embodiment* will be significantly enhanced by *Controller* (**H1A**), *Attached* (**H1B**), *Stretch* (**H1C**) and *Location* (**H1D**).

H2 Users will have different degree of *Embodiment* depending on the interaction mode (*Mode*).

H3 *Embodiment* will be positively influenced by the perceived location (*Proprioception*).

4.5.2 Perceived Location. Hypothesis **H4** suggests that rendering the controller can negatively affect proprioception, whereas having correct multisensory feedback can positively affect it. The possible misalignment between the physical and real location due to the visuo-tactile inconsistency can cause breaks in the embodiment [3]. Conversely, conditions that effectively align visual and tactile stimuli are expected to influence the perceived location accuracy positively [2, 37, 40].

H4 *Proprioception* will be negatively affected by *Controller* (**H4A**), and positively affected by *Attached* (**H4B**), *Stretch* (**H4C**) and *Location* (**H4D**).

4.5.3 Preference. Similarly to the initial set of hypotheses (**H1-H3**), we propose that the different conditions will enhance user preference (**H5**). Although the initial set of hypotheses is based on previous work on the SoE, we hypothesize that increasing the SoE will benefit user preference similarly to the work by Fribourg et al. [9]. Consequently, varying preference levels depend on the interaction mode (**H6**), and increasing embodiment and perceived location will positively impact user preference (**H7**).

H5 *Preference* will be significantly enhanced by *Controller* (**H5A**), *Attached* (**H5B**), *Stretch* (**H5C**) and *Location* (**H5D**).

H6 Users will have different degree of *Preference* depending on the interaction mode (*Mode*).

H7 *Preference* will be positively influenced by *Embodiment* (**H7A**) and *Proprioception* (**H7B**).

4.5.4 Performance. Finally, the last set of hypotheses examines the impact of the different conditions on task performance. Previous studies found that higher performance in some tasks can be achieved when the SoE is high [29, 30, 35]. Therefore, we hypothesize that the same factors that contribute to the SoE will improve performance. We investigate each task individually (**H8-13**). Lastly, we propose that increased embodiment, perceived location, and preference will positively enhance task performance (**H14**).

H8-H10 *Performance_{Cube}* (**H8**), *Performance_{Cannon}* (**H9**) and *Performance_{Painting}* (**H10**) will be significantly enhanced by *Controller* (**H8A-H10A**), *Attached* (**H8B-H10B**), *Stretch* (**H8C-H10C**) and *Location* (**H8D-H10D**).

H11-H13 Users will have different degree of *Performance_{Cube}* (**H11**), *Performance_{Cannon}* (**H12**) and *Performance_{Painting}* (**H13**) depending on the interaction mode (*Mode*).

H14 *Performance* will be positively influenced by *Embodiment* (**H14A**), *Proprioception* (**H14B**) and *Preference* (**H14C**).

5 RESULTS

In this section, we present an overview of the results obtained from the statistical analysis and revise whether the presented hypotheses are substantiated or refuted. For an in-depth discussion and interpretation of these findings, please refer to Section 6.

Shapiro-Wilk tests indicated significant deviations from normality in some instances. As a result, all analyses are carried out using non-parametric tests. To examine the influence of the four conditions—*Controller*, *Attached*, *Stretch*, *Location*—on *Embodiment*, *Preference*, and the performance metrics, we employ Wilcoxon tests. Usage of ANOVA was ruled out due to the insufficient combinations available for the study. However, to mitigate the risk of Type 1 errors, we adjust the p-values using the Bonferroni correction. We also report the Wilcoxon effect size (*r*). We did not observe significant differences based on varying levels of VR experience among participants, therefore, we do not report the results separately.

When studying the differences between interaction modes, we perform a one-way repeated measures ANOVA on ranks (Friedman test) followed by post-hoc tests based on the Wilcoxon test. Similarly, p-values are adjusted with the Bonferroni correction. We present Kendall's *W* effect size for the ANOVA and the Wilcoxon effect size (*r*) for post hoc tests.

Lastly, we leverage linear and binomial mixed-effects models to account for the repeated measures on the same subjects. We employ the binomial model to study *Proprioception* given its binary nature, and linear models to investigate the relationship between *Embodiment*, *Preference*, *Proprioception*, and *Performance*. We scale numerical data when multiple factors are involved for easy comparison.

To gain a deeper understanding of the *Stretch* effect on the avatar, we also measured the maximum arm stretch distance for each user when the *Stretch* condition was active. This distance varies among users, as it is dependent on both the individual's arm length and the virtual avatar's proportions. In our study, we found that the average maximum stretch distance reached 26.3 cm, with a standard deviation of 4.65 cm.

5.1 Embodiment

Table 4 presents the results from the Wilcoxon tests on *Embodiment*. Notably, *Controller* and *Attach* exhibit no significance, and their effect sizes are near zero. On the contrary, both *Stretch* and *Location* yield significant results with moderate effect sizes. Consequently, we dismiss H1A-B and accept H1C-D.

Fig. 6 displays the results of a Friedman test examining the impact of *Mode* on *Embodiment*. It reveals moderate differences among the groups, implying that certain interaction modes result in elevated *Embodiment* levels. Hence, we accept H2. In the post-hoc tests, we observe significant effects when contrasting methods that include *Stretch* with those that do not. However, no distinguishable difference is observed between the two methods that employ *Stretch*.

When we predict *Embodiment* from *Proprioception* using a linear mixed-effects model, we notice a positive effect (Estimate = 2.275, Std. Error = 0.777, $t = 2.928$, $p < .01$). This result aligns with H1C-D, suggesting that when users perceive their hand to be correctly positioned, they experience increased embodiment. Therefore, we accept H3.

Hypothesis	Effect	p	r	
H1A	<i>Controller</i>	0.912	0.007	small
H1B	<i>Attached</i>	0.917	0.009	small
H1C	<i>Stretch</i>	<.001	0.311	moderate
H1D	<i>Location</i>	<.001	0.307	moderate

Table 4: Wilcoxon tests of *Controller*, *Attached*, *Stretch* and *Location* on *Embodiment*. p is the adjusted p-value with Bonferroni correction and r is the Wilcoxon effect size.

5.2 Perceived Hand Location (Proprioception)

In Table 5, the results of the binomial mixed-effects model applied to *Proprioception* are presented. Notably, *Location* does not significantly influence the perceived location, as evidenced by its low odds ratio of 1.809, especially when compared to *Attached* (2.430) and *Stretch* (3.104). It is only *Stretch* that exhibits a significant impact on *Proprioception*. Therefore, we accept H4C and reject H4A, H4B and H4D. The results can further be examined in Fig. 5, where the proportion of participants answering *Yes* to LQ1 exceeded 50% exclusively in the interaction modes incorporating *Stretch*.

5.3 Preference

Table 4 outlines the results from the Wilcoxon tests on *Preference*. All factors significantly influenced *Preference* except for *Attached*. Factors *Stretch* (0.588) and *Location* (0.775) demonstrated large effect sizes. As depicted in Fig. 7, the interaction modes—*FreeController*, *StretchController*, and *StretchHand*—that accurately located the controller were preferred. Fig. 7 also reveals significant differences between interaction modes, particularly between groups that precisely located the controller and those that did not. Therefore, we accept hypotheses H5A, H5C, H5D, and H6, while H5B is rejected.

Upon predicting *Preference* from *Embodiment* (Estimate = 0.459, Std. Error = 0.062, $t = 7.295$, $p < .001$) and *Proprioception*

Hypo.	Effect	Est.	SE	Odds Ratio	z	p
H4A	<i>Controller</i>	0.069	0.295	1.072	0.237	1.000
H4B	<i>Attached</i>	0.888	0.398	2.430	2.232	0.102
H4C	<i>Stretch</i>	1.133	0.319	3.105	3.541	<.01
H4D	<i>Location</i>	0.593	0.296	1.809	2.003	0.180

Table 5: Binomial generalized linear mixed-effects models of *Controller*, *Attached*, *Stretch* and *Location* on *Proprioception*. Est. (Estimate) is the coefficient for the predictor in the logistic model. SE is the standard error of the coefficient estimate. Odds Ratio represents the odds of the user perceiving location when the corresponding effect is set to *yes*. z is the z -value used to determine p . p is the adjusted p-value with Bonferroni correction.

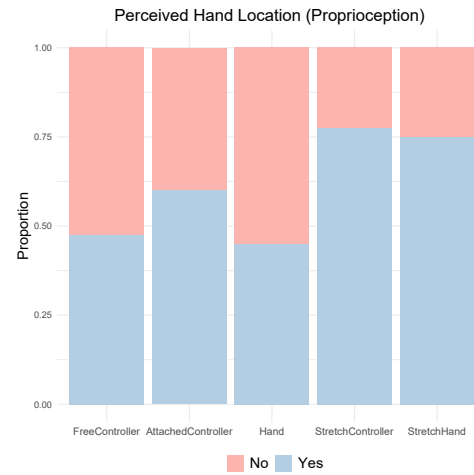


Figure 5: Proportional representation of responses to the question LQ1 *Do you think your hand was in the position of your virtual hand?* (*Proprioception*) across different interaction modes (*Mode*). Each mode is represented by a stacked bar, indicating the proportion of *Yes* (1) and *No* (0) responses.

(Estimate = 0.105, Std. Error = 0.063, $t = 1.673$, $p = 0.096$) using a linear mixed-effects model, both factors seem to exert a positive impact. However, *Embodiment* is the only significant linear predictor of *Preference*. This might be due to *Proprioception* already explaining *Embodiment*, as suggested by hypothesis H3. Thus, we accept H7A and reject H7B.

Hypothesis	Effect	p	r	
H5A	<i>Controller</i>	<.01	0.215	small
H5B	<i>Attached</i>	0.076	0.214	small
H5C	<i>Stretch</i>	<.001	0.588	large
H5D	<i>Location</i>	<.001	0.775	large

Table 6: Wilcoxon tests of *Controller*, *Attached*, *Stretch* and *Location* on *Preference*. p is the adjusted p-value with Bonferroni correction and r is the Wilcoxon effect size.

5.4 Performance

Table 7 present the results from Wilcoxon tests on $Performance_{Cube}$, $Performance_{Cannon}$ and $Performance_{Painting}$. In the context of the *Cube Task*, all factors with the exception of *Controller* significantly impacted performance. The largest effect was seen from *Location* (0.564). For the other tasks, the effects were generally smaller. Notably, *Attached* and *Stretch* significantly influenced performance in the *Cannon Task*, while only *Attached* had a significant impact on the *Painting Task*. As a result, we accept H8B-D, H9B-C, and H10B, and reject H8A, H9A, H9D, H10A, and H10C-D.

Figures 8, 10 and 9 displays the outcomes of a Friedman test evaluating the influence of *Mode* on $Performance_{Cube}$, $Performance_{Cannon}$ and $Performance_{Painting}$. Moderate differences among groups were observed for the *Cube Task*, and smaller differences were noted for the other tasks. Despite these differences being primarily small, significant distinctions were identified, leading us to accept hypotheses H11, H12, and H13.

Finally, in the linear mixed-effects model predicting $Performance$ from *Embodiment* (Estimate = 0.086, Std. Error = 0.076, $t = 1.137$, $p = 0.256$), *Proprioception* (Estimate = 0.139, Std. Error = 0.075, $t = 1.851$, $p = 0.065$), and *Preference* (Estimate = 0.256, Std. Error = 0.077, $t = 3.320$, $p < .01$), all factors appear to contribute positively to $Performance$. However, only *Preference* significantly predicts $Performance$ in a linear manner, with *Proprioception* showing a trend toward positive prediction. Consequently, we accept H14C and reject H14A-B.

Task	Hypothesis	Effect	p	r	
Cube	H8A	<i>Controller</i>	1.0	0.015	small
	H8B	<i>Attached</i>	< .01	0.329	moderate
	H8C	<i>Stretch</i>	< .001	0.364	moderate
	H8D	<i>Location</i>	< .001	0.564	large
Cannon	H9A	<i>Controller</i>	0.169	0.144	small
	H9B	<i>Attached</i>	< .001	0.348	moderate
	H9C	<i>Stretch</i>	< .001	0.269	small
	H9D	<i>Location</i>	1.0	0.009	small
Painting	H10A	<i>Controller</i>	0.051	0.176	small
	H10B	<i>Attached</i>	< .05	0.247	small
	H10C	<i>Stretch</i>	0.181	0.142	small
	H10D	<i>Location</i>	1.0	0.065	small

Table 7: Wilcoxon tests of *Controller*, *Attached*, *Stretch* and *Location* on $Performance_{Cube}$, $Performance_{Cannon}$ and $Performance_{Painting}$. p is the adjusted p-value with Bonferroni correction and r is the Wilcoxon effect size.

6 DISCUSSION AND GUIDELINES

In this section, we aim to understand and interpret the obtained results, while we also try to derive a few guidelines that could help VR designers enhance user embodiment, proprioception, user preference, and/or task performance in their experiences depending on the requirements of their applications.

6.1 Embodiment

When the Sense of Embodiment is essential, do not break body continuity. *Stretch* and *Location* have a statistically significant effect on embodiment. *Stretch* also affects users' proprioception (perceiving their hand in the correct location). These results are also consistent with H3, which states that *Proprioception* positively affects Embodiment. Therefore, we recommend using *Stretch* to guarantee the correct location of end-effectors while providing body continuity, consistent with the results in Dewez et al. [7].

Multisensory coherency with the VR controllers is not necessary to achieve the Sense of Embodiment. We expected to observe higher *Embodiment* when the controller is attached to the hand since it provides multisensory coherence (visual-tactile). However, contrary to our hypothesis, we found a significant moderate effect ($p < .05$, Effect Size = 0.454) on *Embodiment* being higher with the *FreeController* mode than with *AttachedController*. We believe *AttachedController* yielded lower *Embodiment* due to the inconsistency between visual-tactile and proprioceptive feedback during the *Cube Task*, because users could perceive that their real hands were not located where the virtual hands were being rendered. Note that in Section 5.1, we showed that there is a strong correlation between *Proprioception* and *Embodiment*, which could explain our results. However, *FreeController* had higher embodiment despite the conflicting visual-tactile feedback. Other studies in the literature [41] also show that multisensory feedback is not always needed to enhance the SoE. It is also known that certain body illusions can trick our proprioceptors [19]. Further studies are needed to understand how conflicting sensory feedback could be used to modify own-body representation. This would make it possible to fill the gaps caused by the mismatches in dimension and animation.

Maximizing the Sense of Embodiment with *StretchController* and *StretchHand*. Our findings indicate that to enhance the SoE, it is effective to utilize techniques like *StretchController* and *StretchHand*. These methods involve stretching the virtual arms to align with the perceived position of the user's real arms, thereby addressing any mismatches between the user's physical body and their virtual representation while maintaining bodily continuity. Our results also suggest that achieving a high level of SoE does not necessarily require multisensory coherence; thus, stretching the arms is effective across both modes, with (*StretchController*) or without controller rendering (*StretchHand*).

6.2 Perceived hand location

Stretch positively influences the correct perception of hand location. Regarding the perceived hand location or *Proprioception*, most users respond affirmatively to the location of their hands being correct when *Stretch* is enabled since the location of the virtual hand is correctly aligned with the real hand at all times, and body continuity is preserved. Therefore, employing *StretchController* or *StretchHand* is recommended to maximize *Proprioception*. In contrast, while *FreeController* accurately aligns the virtual controller with the real-world counterpart, it negatively impacts *Proprioception*, and thus, should be avoided.

Attached controllers can induce a proprioceptive drift. When *Attached* is enabled, we increase the chances of people believing

that the hand is in the correct position, even though it may not be. Fig. 5 supports this result (over 60 % of participants reported that their hands were perceived in the correct location). The binomial linear model indicated that including *Attached* may substantially increase the probability of users believing their hands were in the right position (Odds Ratio = 2.43). This is known as proprioceptive drift, which is induced by having synchronous visual-tactile feedback [13]. However, our result was not statistically significant. We believe this happens because the *Cube Task* included far-away objects that were hard to reach, whereas, in previous studies, the positions were always within reach. Thus, we recommend attaching controllers for tasks that do not require interaction with objects that are hard to reach. For instance, a dancing application with minimal interaction with the environment will benefit from the *AttachedController* mode.

6.3 Preference

For interaction tasks, provide VR controllers or hand locations as accurately as possible. Both *Stretch* and *Location* are significant and have large effect sizes on *Preference*. But we cannot conclude whether *Stretch* alone affects *Preference* because, in our study, *Stretch* always implies correct *Location*. However, we can conclude that *Location* is essential since even in the *FreeController* mode (in which the controller flies away breaking body continuity and multisensory coherency) *Preference* is high. During the *Cube Task*, participants had to interact with objects that were hard to reach, and some of them commented that they felt frustrated when a position appeared to be physically reachable. Still, their virtual arm could not reach it. This led to participants preferring interaction modes that allowed them to interact correctly. Therefore, from a preference perspective, it appears that *FreeController*, *StretchController*, and *StretchHand* are equally preferred by users.

Maximizing embodiment positively affects preference. As we accepted H7A, creating an experience with self-avatars that enhance embodiment can also increase user preference. Therefore, studies and work focusing on maximizing embodiment will directly positively affect preference.

6.4 Performance

For tasks with hard-to-reach objects, stretching arms provides the best performance without affecting embodiment. In our study, performance differences appeared mostly during the *Cube Task*. In this task, users needed to carefully reach for objects located far away; thus, they could observe the differences between interaction metaphors. Performance is higher when the end-effector is in the correct location, which occurs when we stretch the arm or use the flying controller. Similarly to *Preference*, modes *FreeController*, *StretchController*, and *StretchHand* would equally perform for pick-and-place tasks.

The *Cannon task* requires very rapid movements, in which the user was mainly focused on the next ball, that they could hardly observe differences, but even then, performance was worse for the *FreeController* mode.

For the *Painting Task*, the range of movements was mostly limited to a natural arm position, which did not often trigger the arm stretching or the flying controller. However, the few times these

differences were observed led to users performing worse with the *FreeController* mode.

Therefore, for tasks that require accurate interaction with virtual objects, we recommend virtually stretching the arm to allow for correct end-effector positioning while respecting visual-tactile and proprioceptive coherence. For applications that require manipulation of objects within easy reach or that the interaction requires rapid movements without careful manipulation, the stretching technique would also be valid but not necessary. So, in these cases, having attached controllers would suffice.

7 LIMITATIONS

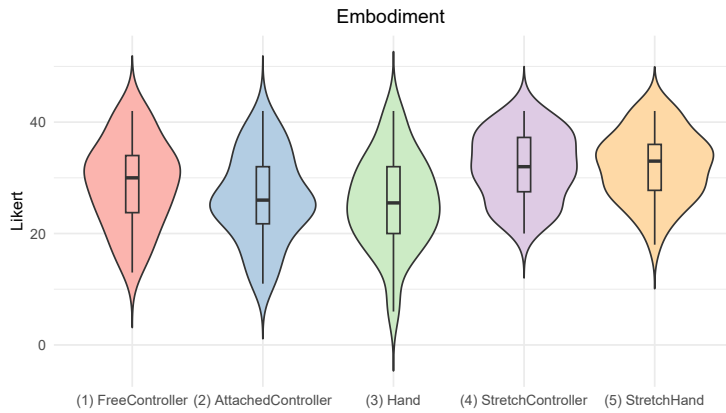
This section outlines the primary limitations identified in our study. Firstly, the results are derived from a specific set of tasks, primarily centered around interaction tasks designed to mimic everyday activities in VR. This focus means our findings might not fully represent the variety of experiences in VR applications. For instance, in scenarios without active interaction, users might be less likely to notice the incorrect alignment of virtual and real hands, as suggested by Ponton et al. [32].

The generalizability of our results is also potentially limited by the demographics of our participant sample. Most participants were university students with limited VR experience. While embodiment questionnaires are a widely accepted tool, some participants might have faced challenges understanding and accurately responding to them. We minimized confounding factors, yet certain technical aspects, such as pose animation and the designs of avatars and controllers, could have influenced the outcomes. Additionally, we believe that the skinning of the avatar when stretching the arms can significantly affect the results.

Finally, our study did not investigate every possible combination of factors like *Stretch* and *Location* or the case where the virtual arm could be longer than the real arm. To effectively separate the effects of *Stretch* and *Location*, we would have needed to include an interaction mode where the arm is stretched but not precisely located in the correct position. For instance, Kiltani et al. [20] studies the SoE based on different arm lengths. However, such a mode does not align well with practical VR scenarios, leading us to exclude it from our study. This decision, while rational from a practical standpoint, does limit the comprehensiveness of our findings regarding these factors. Note that if the virtual arm were longer than the real hand, the result would be that the IK would bend at the elbow to reach the end-effector. This situation does not prevent the user from reaching objects, it simply introduces a mismatch in the pose of the arm [32], but not in the position of the end-effector. Our *Stretch* was limited to the values needed to fill the mismatch between the virtual and real hand, we have not investigated the tolerable mismatch threshold and its impact on embodiment.

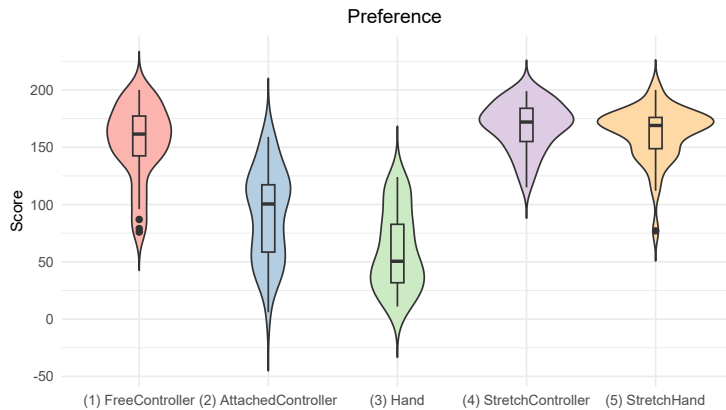
8 CONCLUSIONS AND FUTURE WORK

In this paper, we have studied 5 interaction metaphors for animated full-body avatars. We have focused our efforts on the simulation and rendering of arms, hands, and controllers. Our results suggest that selecting the best interaction metaphor strongly depends on the type of task that the user needs to perform. However, results on all embodiment, performance, preference, and proprioception suggest



Effect	p	Effect Size	
<i>Mode</i>	<.001	0.314	moderate
1 - 2	<.05	0.454	moderate
1 - 3	0.06	0.441	moderate
1 - 4	<.01	0.554	large
1 - 5	<.05	0.483	moderate
2 - 3	1.0	0.139	small
2 - 4	<.001	0.702	large
2 - 5	<.001	0.716	large
3 - 4	<.001	0.749	large
3 - 5	<.001	0.754	large
4 - 5	1.0	0.021	small

Figure 6: Hypothesis H2. One-way repeated measures ANOVA on ranks (Friedman test) of *Mode* on *Embodiment* and the corresponding post-hoc tests (Wilcoxon signed-rank test). p is the adjusted p -value with Bonferroni correction. Effect Size is Kendall's W for the Friedman test and the r value for the post-hoc tests.



Effect	p	Effect Size	
<i>Mode</i>	<.001	0.764	large
1 - 2	<.001	0.860	large
1 - 3	<.001	0.871	large
1 - 4	0.589	0.300	small
1 - 5	1.0	0.174	small
2 - 3	<.001	0.798	large
2 - 4	<.001	0.871	large
2 - 5	<.001	0.850	large
3 - 4	<.001	0.871	large
3 - 5	<.001	0.871	large
4 - 5	0.997	0.267	small

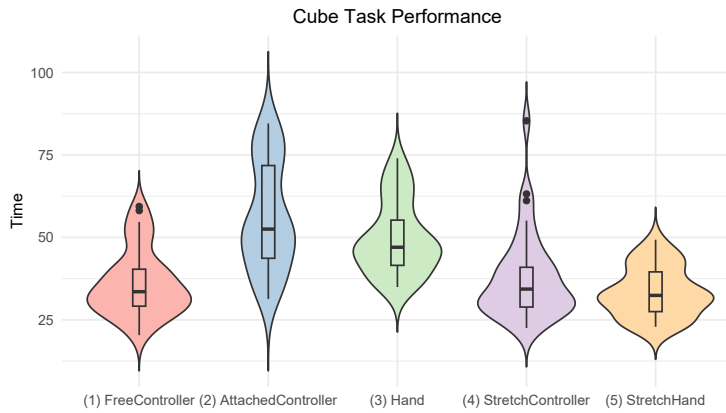
Figure 7: Hypothesis H6. One-way repeated measures ANOVA on ranks (Friedman test) of *Mode* on *Preference* and the corresponding post-hoc tests (Wilcoxon signed-rank test). p is the adjusted p -value with Bonferroni correction. Effect Size is Kendall's W for the Friedman test and the r value for the post-hoc tests.

that the best interaction metaphor is to have stretch arms that can allow the end-effectors to be in the correct position while respecting body continuity. Rendering the controllers does not appear to be relevant, although attaching the controllers to the hand may induce a proprioceptive drift due to the consistent visual-tactile feedback.

In future work, we would like to study interaction metaphors when performing collaborative tasks, for example in cases where two participants need to pass a virtual object or carry something together. We would also like to experiment with how interaction metaphors could be used for the inclusiveness of people with mobility limitations, by improving their interaction in VR beyond their possibilities in the real world. For example, extending the stretch interaction mode to help them reach for objects that are slightly beyond their physical reach, without breaking embodiment or body continuity.

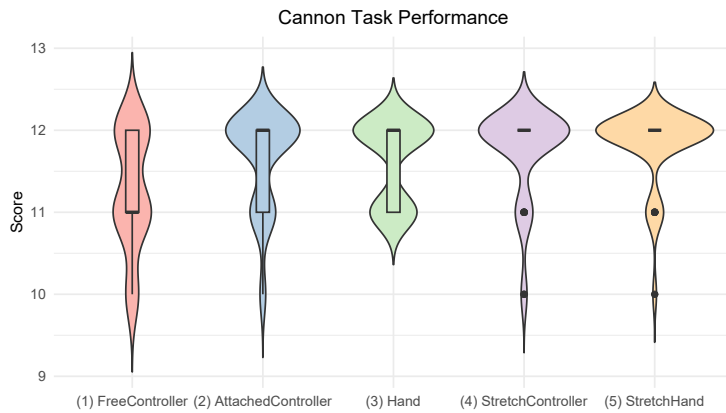
ACKNOWLEDGMENTS

This work has received funding from the European Union's Horizon 2020 research and innovation programme under HORIZON-CL4-2022-HUMAN-01 grant agreement No 101093159 (XR4ED), and from MCIN/AEI/10.13039/501100011033/FEDER "A way to make Europe", UE (PID2021-122136OB-C21). Jose Luis Ponton was also funded by the Spanish Ministry of Universities (FPU21/01927).



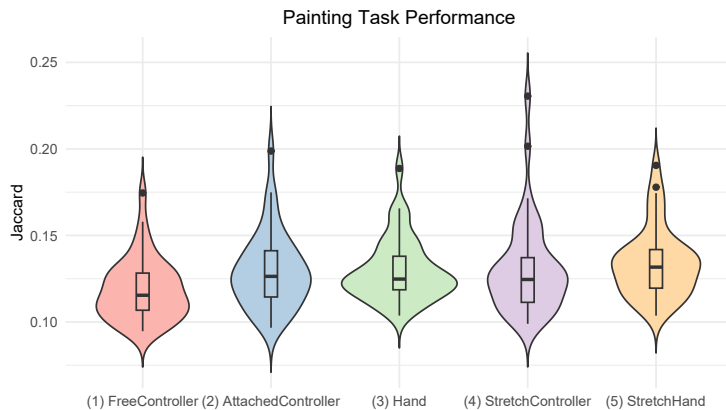
Effect	p	Effect Size	
Mode	<.001	0.403	moderate
1 - 2	<.001	0.771	large
1 - 3	<.001	0.724	large
1 - 4	1.0	0.167	small
1 - 5	1.0	0.093	small
2 - 3	1.0	0.244	small
2 - 4	<.01	0.552	large
2 - 5	<.001	0.667	large
3 - 4	<.05	0.482	moderate
3 - 5	<.001	0.639	large
4 - 5	1.0	0.199	small

Figure 8: Hypothesis H11. One-way repeated measures ANOVA on ranks (Friedman test) of Mode on Performance_{Cube} and the corresponding post-hoc tests (Wilcoxon signed-rank test). p is the adjusted p-value with Bonferroni correction. Effect Size is Kendall's W for the Friedman test and the r value for the post-hoc tests.



Effect	p	Effect Size	
Mode	<.001	0.196	small
1 - 2	0.117	0.402	moderate
1 - 3	0.086	0.446	moderate
1 - 4	<.01	0.629	large
1 - 5	<.001	0.705	large
2 - 3	1.0	0.012	small
2 - 4	1.0	0.207	small
2 - 5	0.510	0.298	small
3 - 4	1.0	0.154	small
3 - 5	1.0	0.283	small
4 - 5	1.0	0.094	small

Figure 9: Hypothesis H12. One-way repeated measures ANOVA on ranks (Friedman test) of Mode on Performance_{Cannon} and the corresponding post-hoc tests (Wilcoxon signed-rank test). p is the adjusted p-value with Bonferroni correction. Effect Size is Kendall's W for the Friedman test and the r value for the post-hoc tests.



Effect	p	Effect Size	
Mode	<.01	0.104	small
1 - 2	<.05	0.469	moderate
1 - 3	<.05	0.448	moderate
1 - 4	0.076	0.416	moderate
1 - 5	<.001	0.648	large
2 - 3	1.0	0.040	small
2 - 4	1.0	0.034	small
2 - 5	1.0	0.193	small
3 - 4	1.0	0.065	small
3 - 5	1.0	0.212	small
4 - 5	1.0	0.197	small

Figure 10: Hypothesis H13. One-way repeated measures ANOVA on ranks (Friedman test) of Mode on Performance_{painting} and the corresponding post-hoc tests (Wilcoxon signed-rank test). p is the adjusted p-value with Bonferroni correction. Effect Size is Kendall's W for the Friedman test and the r value for the post-hoc tests.

REFERENCES

- [1] Ferran Argelaguet, Ludovic Hoyet, Michael Trico, and Anatole Lecuyer. 2016. The Role of Interaction in Virtual Embodiment: Effects of the Virtual Hand Representation. In *2016 IEEE Virtual Reality (VR)*. 3–10. <https://doi.org/10.1109/VR.2016.7504682>
- [2] Matthew Botvinick and Jonathan Cohen. 1998. Rubber Hands ‘Feel’ Touch That Eyes See. *Nature* 391, 6669 (Feb. 1998), 756–756. <https://doi.org/10.1038/35784>
- [3] E. Burns, S. Razzaque, A.T. Panter, M.C. Whitton, M.R. McCallus, and F.P. Brooks. 2005. The hand is slower than the eye: a quantitative exploration of visual dominance over proprioception. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005*. 3–10. <https://doi.org/10.1109/VR.2005.1492747>
- [4] Giuseppe Caggianese, Luigi Gallo, and Pietro Neroni. 2019. The Vive Controllers vs. Leap Motion for Interactions in Virtual Environments: A Comparative Evaluation. In *Intelligent Interactive Multimedia Systems and Services (Smart Innovation, Systems and Technologies)*, Giuseppe De Pietro, Luigi Gallo, Robert J. Howlett, Lakhmi C. Jain, and Ljubo Vlacic (Eds.). Springer International Publishing, Cham, 24–33. https://doi.org/10.1007/978-3-319-92231-7_3
- [5] Lucio Tommaso De Paolis and Valerio De Luca. 2020. The Impact of the Input Interface in a Virtual Environment: The Vive Controller and the Myo Armband. *Virtual Reality* 24, 3 (Sept. 2020), 483–502. <https://doi.org/10.1007/s10055-019-00409-6>
- [6] Lucio Tommaso De Paolis and Valerio De Luca. 2022. The Effects of Touchless Interaction on Usability and Sense of Presence in a Virtual Environment. *Virtual Reality* 26, 4 (Dec. 2022), 1551–1571. <https://doi.org/10.1007/s10055-022-00647-1>
- [7] Diane Dewez, Ludovic Hoyet, Anatole Lécuyer, and Ferran Argelaguet Sanz. 2021. Towards “Avatar-Friendly” 3D Manipulation Techniques: Bridging the Gap Between Sense of Embodiment and Interaction in Virtual Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–14. <https://doi.org/10.1145/3411764.3445379>
- [8] Tiare Feuchtnr and Jörg Müller. 2017. Extending the Body for Interaction with Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI ’17)*. Association for Computing Machinery, New York, NY, USA, 5145–5157. <https://doi.org/10.1145/3025453.3025689>
- [9] Rebecca Fribourg, Ferran Argelaguet, Anatole Lécuyer, and Ludovic Hoyet. 2020. 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 (May 2020), 2062–2072. <https://doi.org/10.1109/TVCG.2020.2973077>
- [10] Meiyuzi Gao and Deborah A. Boehm-Davis. 2023. Development of a Customizable Interactions Questionnaire (CIQ) for Evaluating Interactions with Objects in Augmented/Virtual Reality. *Virtual Reality* 27, 2 (June 2023), 699–716. <https://doi.org/10.1007/s10055-022-00678-8>
- [11] Guilherme Gonçalves, Miguel Melo, Luis Barbosa, José Vasconcelos-Raposo, and Maximino Bessa. 2022. Evaluation of the Impact of Different Levels of Self-Representation and Body Tracking on the Sense of Presence and Embodiment in Immersive VR. *Virtual Reality* 26, 1 (March 2022), 1–14. <https://doi.org/10.1007/s10055-021-00530-5>
- [12] Eric J Gonzalez and Sean Follmer. 2023. Sensorimotor Simulation of Redirected Reaching Using Stochastic Optimal Feedback Control. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (Hamburg, Germany) (CHI ’23)*. Association for Computing Machinery, New York, NY, USA, Article 776, 17 pages. <https://doi.org/10.1145/3544548.3580767>
- [13] Mar Gonzalez-Franco, Brian Cohn, Eyal Ofek, Dalila Burin, and Antonella Maselli. 2020. The Self-Avatar Follower Effect in Virtual Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 18–25. <https://doi.org/10.1109/VR46266.2020.00019>
- [14] Mar Gonzalez-Franco and Tabitha C. Peck. 2018. Avatar Embodiment. Towards a Standardized Questionnaire. *Frontiers in Robotics and AI* 5 (2018). <https://doi.org/10.3389/frobt.2018.00074>
- [15] Elisa Gusai, Chiara Bassano, Fabio Solari, and Manuela Chessa. 2017. Interaction in an Immersive Collaborative Virtual Reality Environment: A Comparison Between Leap Motion and HTC Controllers. In *New Trends in Image Analysis and Processing – ICIAP 2017 (Lecture Notes in Computer Science)*, Sebastiano Battiato, Giovanni Maria Farinella, Marco Leo, and Giovanni Gallo (Eds.). Springer International Publishing, Cham, 290–300. https://doi.org/10.1007/978-3-319-70742-6_27
- [16] Ludovic Hoyet, Anatole Le, Diane Dewez, Ferran Argelaguet, et al. 2023. To Stick or Not to Stick? Studying the Impact of Offset Recovery Techniques During Mid-Air Interactions. *IEEE Transactions on Visualization and Computer Graphics* (2023). <https://doi.org/10.1109/TVCG.2023.3295209>
- [17] M.-Carmen Juan, Julen Elexpuru, Paulo Dias, Beatriz Sousa Santos, and Paula Amorim. 2023. Immersive Virtual Reality for Upper Limb Rehabilitation: Comparing Hand and Controller Interaction. *Virtual Reality* 27, 2 (June 2023), 1157–1171. <https://doi.org/10.1007/s10055-022-00722-7>
- [18] Konstantina Kilteni, Raphaela Groten, and Mel Slater. 2012. The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators and Virtual Environments* 21, 4 (Nov. 2012), 373–387. https://doi.org/10.1162/PRES_a_00124
- [19] Konstantina Kilteni, Antonella Maselli, Konrad P Kording, and Mel Slater. 2015. Over my fake body: body ownership illusions for studying the multisensory basis of own-body perception. *Frontiers in human neuroscience* 9 (2015), 141. <https://doi.org/10.3389/fnhum.2015.00141>
- [20] Konstantina Kilteni, Jean-Marie Normand, Maria V. Sanchez-Vives, and Mel Slater. 2012. Extending Body Space in Immersive Virtual Reality: A Very Long Arm Illusion. *PLoS One* 7, 7 (July 2012), e40867. <https://doi.org/10.1371/journal.pone.0040867>
- [21] Silvia Erika Kober, Volker Settgast, Marlies Brunnhofer, Ursula Augsdörfer, and Guilherme Wood. 2022. Move Your Virtual Body: Differences and Similarities in Brain Activation Patterns during Hand Movements in Real World and Virtual Reality. *Virtual Reality* 26, 2 (June 2022), 501–511. <https://doi.org/10.1007/s10055-021-00588-1>
- [22] Lorraine Lin and Sophie Jörg. 2016. Need a Hand? How Appearance Affects the Virtual Hand Illusion. In *Proceedings of the ACM Symposium on Applied Perception (SAP ’16)*. Association for Computing Machinery, New York, NY, USA, 69–76. <https://doi.org/10.1145/2931002.2931006>
- [23] Joan Llobera, Alejandro Beacco, Ramon Oliva, Gizem Senel, Domna Banakou, and Mel Slater. 2021. Evaluating Participant Responses to a Virtual Reality Experience Using Reinforcement Learning. *Royal Society Open Science* 8, 210537 (2021). <https://doi.org/10.1098/rsos.210537>
- [24] Christos Lougiakis, Akrivi Katifori, Maria Roussou, and Ioannis-Panagiotis Ioannidis. 2020. Effects of Virtual Hand Representation on Interaction and Embodiment in HMD-based Virtual Environments Using Controllers. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 510–518. <https://doi.org/10.1109/VR46266.2020.00072>
- [25] Antonella Maselli and Mel Slater. 2013. The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience* 7 (2013). <https://doi.org/10.3389/fnhum.2013.00083>
- [26] Roberto A. Montano Murillo, Sriram Subramanian, and Diego Martinez Plasencia. 2017. Erg-O: Ergonomic Optimization of Immersive Virtual Environments. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (Québec City, QC, Canada) (UIST ’17)*. Association for Computing Machinery, New York, NY, USA, 759–771. <https://doi.org/10.1145/3126594.3126605>
- [27] Solène Neyret, Xavi Navarro, Alejandro Beacco, Ramon Oliva, Pierre Bourdin, Jose Valenzuela, Itxaso Barberia, and Mel Slater. 2020. An Embodied Perspective as a Victim of Sexual Harassment in Virtual Reality Reduces Action Conformity in a Later Milgram Obedience Scenario. *Scientific Reports* 10, 6207 (2020). <https://doi.org/10.1038/s41598-020-62932-w>
- [28] Ramon Oliva, Alejandro Beacco, Xavi Navarro, and Mel Slater. 2022. QuickVR: A Standard Library for Virtual Embodiment in Unity. *Frontiers in Virtual Reality* 3 (2022). <https://doi.org/10.3389/frvir.2022.937191>
- [29] Ye Pan and Anthony Steed. 2019. Avatar Type Affects Performance of Cognitive Tasks in Virtual Reality. In *Proc. VRST. ACM*, New York, 1–4. <https://doi.org/10.1145/3359996.3364270>
- [30] Ye Pan and Anthony Steed. 2019. How Foot Tracking Matters: The Impact of an Animated Self-Avatar on Interaction, Embodiment and Presence in Shared Virtual Environments. *Frontiers in Robotics and AI* 6 (Oct. 2019). <https://doi.org/10.3389/frobt.2019.00104>
- [31] Tabitha C. Peck and Mar Gonzalez-Franco. 2021. Avatar Embodiment. A Standardized Questionnaire. *Frontiers in Virtual Reality* 1 (2021). <https://doi.org/10.3389/frvir.2020.575943>
- [32] Jose Luis Ponton, Victor Ceballos, Lesly Acosta, Alejandro Rios, Eva Monclus, and Nuria Pelechano. 2023. Fitted avatars: automatic skeleton adjustment for self-avatars in virtual reality. *Virtual Reality* 27 (July 2023), 2541–2560. <https://doi.org/10.1007/s10055-023-00821-z>
- [33] Jose Luis Ponton, Eva Monclus, and Nuria Pelechano. 2022. AvatarGo: Plug and Play Self-Avatars for VR. In *Eurographics 2022 - Short Papers*. The Eurographics Association, Reims, France, 77–80. <https://doi.org/10.2312/egs20221037>
- [34] Aylene Ricca, Amine Chellali, and Samir Otrane. 2021. The Influence of Hand Visualization in Tool-Based Motor-Skills Training, a Longitudinal Study. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 103–112. <https://doi.org/10.1109/VR50410.2021.00031>
- [35] Brian Ries, Victoria Interrante, Michael Kaeding, and Lane Phillips. 2009. Analyzing the Effect of a Virtual Avatar’s Geometric and Motion Fidelity on Ego-Centric Spatial Perception in Immersive Virtual Environments. In *Proc. VRST. ACM*, New York, 59–66. <https://doi.org/10.1145/1643928.1643943>
- [36] Michael Rietzler, Florian Geiselhart, Jan Gugenheimer, and Enrico Rukzio. 2018. Breaking the Tracking: Enabling Weight Perception Using Perceivable Tracking Offsets. 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/10.1145/3173574.3173702>
- [37] Maria V. Sanchez-Vives, Bernhard Spanlang, Antonio Frisoli, Massimo Bergamasco, and Mel Slater. 2010. Virtual Hand Illusion Induced by Visuomotor Correlations. *PLOS ONE* 5, 4 (April 2010), e10381. <https://doi.org/10.1371/journal.pone.0010381>
- [38] Mel Slater, Domna Banakou, Alejandro Beacco, Jaime Gallego Vila, Francisco Macia-Varela, and Ramon Oliva. 2022. A Separate Reality: An Update on Place

- Illusion and Plausibility in Virtual Reality. *Frontiers in Virtual Reality* 3:914392 (2022). <https://doi.org/10.3389/frvir.2022.914392>
- [39] Mel Slater, Xavi Navarro, Jose Valenzuela, Ramon Oliva, Alejandro Beacco, Jacob Thorn, and Zillah Watson. 2018. Virtually Being Lenin Enhances Presence and Engagement in a Scene from the Russian Revolution. *Frontiers in Robotics and AI* 5:91. (2018). <https://doi.org/10.3389/frobt.2018.00091>
- [40] Mel Slater, Daniel Pérez Marcos, Henrik Ehrsson, and Maria Sanchez-Vives. 2008. Towards a Digital Body: The Virtual Arm Illusion. *Frontiers in Human Neuroscience* 2 (2008). <https://doi.org/10.3389/neuro.09.006.2008>
- [41] Anthony Steed and Vit Drga. 2023. Tomato Presence: Virtual Hand Ownership with a Disappearing Hand. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 649–650. <https://doi.org/10.1109/VRW58643.2023.00167>
- [42] Yujie Tao, Cheng Yao Wang, Andrew D Wilson, Eyal Ofek, and Mar Gonzalez-Franco. 2023. Embodying Physics-Aware Avatars in Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (*CHI '23*). Association for Computing Machinery, New York, NY, USA, Article 254, 15 pages. <https://doi.org/10.1145/3544548.3580979>
- [43] Tanh Quang Tran, HyunJu Shin, Wolfgang Stuerzlinger, and JungHyun Han. 2017. Effects of Virtual Arm Representations on Interaction in Virtual Environments. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (VRST '17)*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3139131.3139149>
- [44] Maurizio Vergari, Tanja Kojić, Francesco Vona, Franca Garzotto, Sebastian Möller, and Jan-Niklas Voigt-Antons. 2021. Influence of Interactivity and Social Environments on User Experience and Social Acceptability in Virtual Reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 695–704. <https://doi.org/10.1109/VR50410.2021.00096>
- [45] Haoran Yun, Jose Luis Ponton, Carlos Andujar, and Nuria Pelechano. 2023. Animation Fidelity in Self-Avatars: Impact on User Performance and Sense of Agency. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. IEEE, 286–296. <https://doi.org/10.1109/VR55154.2023.00044>