Dual Body Bimanual Coordination in Immersive Environments

James Smith
University of California, Berkeley
Berkeley, CA, USA
james.smith@berkeley.edu

Xinyun Cao
University of California, Berkeley
Berkeley, CA, USA
xinyun_c@berkeley.edu

Adolfo G. Ramirez-Aristizabal
Accenture
San Francisco, CA, USA
adolfo.ramirez@accenture.com

Bjoern Hartmann
University of California, Berkeley
Berkeley, CA, USA
bjoern@berkeley.edu

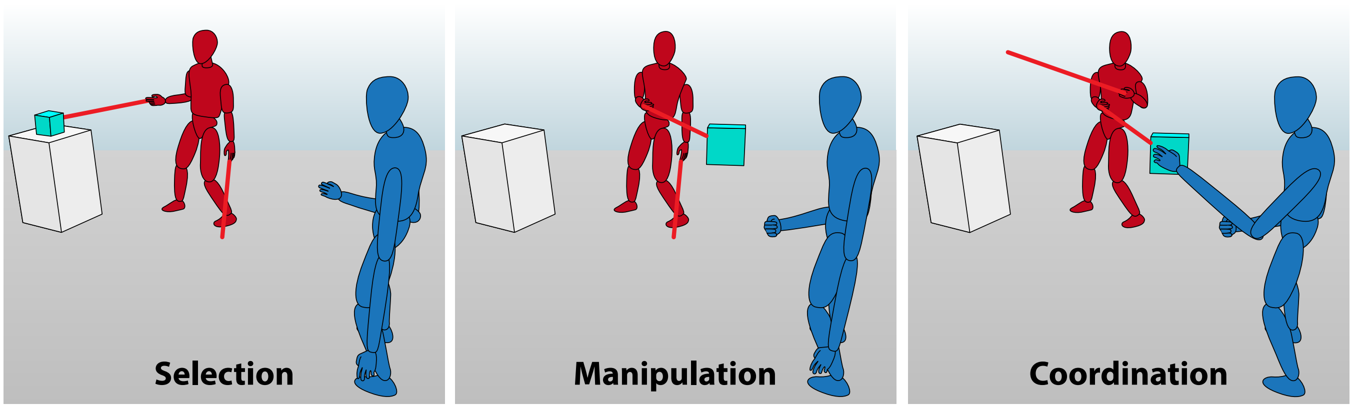


Figure 1: We present Dual Body Bimanual Coordination, an empirical study where users control and interact with the world through two bodies in virtual reality simultaneously. Users select and manipulate objects to perform a coordinated handoff between two bodies under the control of a single user.

ABSTRACT

A common way to enable immersion in VR is to render a virtual body that mirrors the user's physical movements. VR allows us to design interaction schemes that go beyond direct avatar embodiments. In particular, there is a growing body of literature investigating the simultaneous control of multiple bodies in VR. We contribute to this literature by investigating the important case where multiple bodies perform a coordinated interaction with each other. Such actions directly question what kind of embodiment users experience. Concretely, we investigate people's abilities to perform coordinated bimanual selection and handoff tasks between a first-person and third-person body through a user study with 19 participants. Results provide quantitative & qualitative evidence for people's ability to perform complex coordinated tasks through two bodies. Furthermore we characterize participant performance in different task and interaction configurations, summarize the strategies they employed, and discuss qualities of user's proprioception.

DIS '23, July 10–14, 2023, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-9893-0/23/07.
https://doi.org/10.1145/3563657.3596082

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; **Empirical studies in HCI**; *HCI theory, concepts and models.*

KEYWORDS

Virtual Reality, Embodiment, Bimanual Coordination, Virtual Body Schema

ACM Reference Format:
James Smith, Xinyun Cao, Adolfo G. Ramirez-Aristizabal, and Bjoern Hartmann. 2023. Dual Body Bimanual Coordination in Immersive Environments. In *Designing Interactive Systems Conference (DIS '23), July 10–14, 2023, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3563657.3596082

1 INTRODUCTION

Understanding the foundations of embodied interaction has long been a focus of HCI researchers aiming to design better interactive systems and interaction techniques. This is true both in physical/tangible user interfaces [26] as well as in immersive virtual environments [50]. Theories of embodiment can inform the design process itself [11], as well as the artifacts being designed, from interaction techniques in software [26] to human-robot interaction [32].

Cognitive mechanics in coordination and the embodiment of perception-action has been significantly studied in the Cognitive Sciences; with related subfields such as Embodied Cognition and

Complex & Dynamical Systems [29, 41]. Of particular interest is work associated with bimanual interaction and coordination [24, 25]. Such work raises deep questions about the relationship between our minds, bodies, and environment. Advancements in technology and interaction design extend the scope of these questions into immersive virtual environments (VE), allowing us to empirically test user control performance and experience.

One appealing aspect of VR is that researchers can experiment with interactions that go beyond the limits of our physical body's ecological affordances [8]. Prior work exemplifies such an endeavor in the service of selection and manipulation of objects. Techniques such as Go-Go [44] apply proportional gains to the user's hand motions allowing users to reach beyond where their body normally could. Other techniques such as Ninja Hands [48] suggest that having multiple possible locations of interaction is valuable in allowing users to reach far while requiring minimal physical motion from the user.

There have been a number of recent projects exploring the concept of controlling multiple virtual bodies simultaneously in VR [27, 39, 55]. These projects all contain different studies on synchronized control of multiple bodies in VR. Synchronized refers to the fact that the motion of the user is replicated across all bodies that they control, and that all bodies do the same thing. The tasks in these experiments require that users take actions through single bodies and don't explore how the bodies can be coordinated to complete tasks. People's physical capabilities in bimanual control have been well studied [24, 25], and for this reason we investigate bimanual tasks between the two virtual bodies controlled by a single user.

The motivation of this work is to gain a fundamental understanding of whether or not this type of interaction is possible for humans to perform. Given such an understanding, it may be possible to design novel interaction techniques in the future that are based on an understanding of the capabilities and limitations of multi body interaction. It is possible that, for certain tasks, interaction with VEs through multiple bodies would be preferable to single body interaction. In addition, it might be possible to extend multi-body control into the physical world (e.,g., as in [55]), allowing us to multiply our physical actions by controlling multiple robots simultaneously.

In this study we attempt to explore people's abilities to perform coordination interactions between two bodies controlled by one person (see Figure 1). We designed the VE so that users would be required to pick up an object with one body and hand it off to the other body, forcing participants to perform a coordinated action. The motions of the bodies in our VE are synchronized and the user is able to interact with the VE through the bodies simultaneously. The user performs this hand off many times while we vary different variables relating to body positioning and task difficulty.

The user's point of view is always from the perspective of one of the two bodies, so while in the VE they have a first person body and a third person body. We ran two versions of the experiment. In each version, the point of view changes, so in one experiment their first person body is doing the handing off, and in the other version of the experiment their first person body is doing the receiving.

Our study contributes a characterization of user's performance in dual body bimanual coordination. We show that there is a task performance drop off beyond certain angles of rotation of the third person body away from the user's physical body orientation. We also classify user's strategies employed when solving this task, and show that over time user strategy choice trends toward higher performing strategies. Qualitative results show that users did not report feeling embodied in the third person body, however this did not prevent them from successfully acting through the body. Finally, we provide qualitative observations that describe user's lack of awareness of the virtual limbs not involved in the hand off.

2 RELATED WORK

2.1 The Body in Virtual Reality

People's perception has the remarkable ability to extend bodily ownership to coupled environmental stimuli, in which the visual sense is of relevance to VR. This is exemplified by Botvinick's [5] *Rubber hands* experiment, where the simultaneous touch of a physical and artificial hand induced a sense of ownership over the artificial. This idea has been extended into VR by exploring the concept of embodiment, or the body illusion [52] - when we are in a VR environment we gain a sense of ownership over the virtual body presented to us. There have been multiple projects [43, 53] showing that users in virtual reality systems can even perceive another body as their own, as long as they adopt a first person perspective of it. Self-location is a very related concept, which is the position in the environment that people feel like they are located. Furlanetto et al. [14] discuss the idea of mental bilocation - that our mental self-location can be in multiple places at once. They discuss several studies on body illusion during synchronized stroking experiments in VR [30, 31]. Their results show that users do self-localize toward their virtual bodies.

This topic of body illusion has also been explored for third person body ownership. Maselli and Slater [36] show that a user's sense of self location was affected when perceiving a virtual third person representation of themselves. Gorisse [17] explores the ability for users to control a third person body in VR, which showed that participants felt a sense of agency over a third person body despite them not embodying it. Gonzales-Franco et al. [16] present a study where they are able to provide a body ownership illusion of a third person body in VR through a mirror illusion. Nakul et al. [40] present a study measuring the sense of self location with a third person avatar. They show that users do self locate with the third person body under mental imagery tasks. We believe that the sense of agency and self location with third person bodies is a highly valuable avenue for designing interactions utilizing a third person avatars in virtual environments.

Jaron Lanier calls the idea of controlling virtual bodies with morphologies very different than our own "Homuncular Flexibility" [28]. Won, Stevenson and Lanier developed the idea further in [57] where they show the ability of users to control a third virtual arm to accomplish tasks more efficiently in VR. We take inspiration from this idea, and believe controlling multiple humanoid bodies can be learned in a similar way.

2.2 Multiple Bodies in VR

There have also been a number of projects investigating simultaneous control and interaction through multiple bodies. Ninja Hands [48] proposes a technique that presents an array of hands mapped to the user's hand, and interaction can happen through any of them. In OVRLap [49], Schjerlund proposes simultaneous interaction from multiple locations through a single first person perspective. The different reference frames are multiplexed by rendering other points of view to the user in the first person. We also study the ability for users to interact through multiple bodies, but we specifically look at the case where the user is controlling two virtual bodies and accessing a first and third body perspective through a single view.

In MultiSoma [39], users control multiple synchronized virtual bodies located in different locations to perform a selection task, and are shown each body's point of view simultaneously using a view splitting interface. Parallel Adaptation [56] provides a study where users sequentially control multiple bodies in first and third person conditions and are able to experience different motor adaptations between the bodies. Heydrich et al. [20] have shown that users are able to self-identify with two bodies simultaneously in virtual reality. We also explore the concepts of first and third person control, but one notable difference between all of this work and our own is that we present a task to the user where they must coordinate an interaction between their multiple bodies.

2.3 Bimanual Interaction and Proprioception

Bimanual interaction in computing interfaces has long been a topic of inquiry. Buxton and Myers [6] argue that bimanual input is more natural, and their studies show that bimanual input can be used as a way to increase throughput. Guiard's Kinematic Chain [19] proposes a model of bimanual action where the hands are abstracted as motors and can be assembled in a serial linkage, forming a kinematic chain. This model has informed a number of experiments studying bimanual cooperation[2, 23]. Hinckley [21] finds that users rely much more on visual feedback when performing alignment tasks than when doing the tasks bimanually. They find that using both hands allows users to find a sense of space between two reference frames. Other work by Hinckley [22] includes a bimanual interface for neurosurgical applications. They make the argument that parallel bimanual work, an often default interaction scheme, does not always save time due to the hierarchical control dynamics of bimanual manipulation. Therefore, heterogeneous interaction control schemes, such as the one used in our study, are of focus to test the capacity of user performance and coordination.

One of the benefits of VR interfaces is the engagement of our proprioceptive senses. Work on proprioception seeks to explain how this engagement between our perception and action is represented in our brain [18, 33, 34] to outline the phenomena of a peripersonal space awareness surrounding our body. Mine et al. [38] explore how our proprioceptive sense can help us manipulate objects in virtual spaces. Results from their studies show that users had a preference and performed better when grabbing objects in reference spaces relative to their arms. We further explore people's performance in selection and manipulation tasks in different reference frames in our study.

3 STUDY CONCEPTS

The goal of this study is to further our understanding of how users can control multiple bodies in virtual environments. In particular we are interested in how people are able to perform coordinated actions through multiple bodies. In our VE, users will be controlling two bodies simultaneously, and thus have two relationships to the bodies. The user "inhabits" a **first person** body, from which their point of view of the scene originates from. They also control a **third person** body. Each of the bodies exists in it's own **reference frame**, the coordinate system in which it is located. In our study we manipulate the position and rotation of these reference frames. This requires the user to learn a **mapping transform** [4] from their physical frame of reference into each body's reference frame, in order to control the bodies.

People naturally have the ability to perform coordinated bimanual actions, as well as have the ability to perform coordinated actions with other people. We combine these two concepts by giving the user the task of picking up an object with one of the bodies and handing it to the other body. Both hands are needed for this task - one to pick up the object, and the other to receive it, making this a **bimanual coordination task**, where the user's arms are operating through different mapping transforms. The user's **body schema** (mental conception of their body's form and function) changes as users choose interact through these bodies.

How do we decide which bodies to give the users, and where should these bodies be located such that they facilitate a handoff? We provide **different interaction affordances** for each body, which has two effects. First, it allows us to define a model for how to limit the number of possible body position configurations to study, while still providing the user with a task that can be completed (See Figure 3) by restricting the number of viable body positions. Secondly, we are able to vary the parameters of this model as the parameters to the study, to explore how positioning can affect performance.

4 METHODS

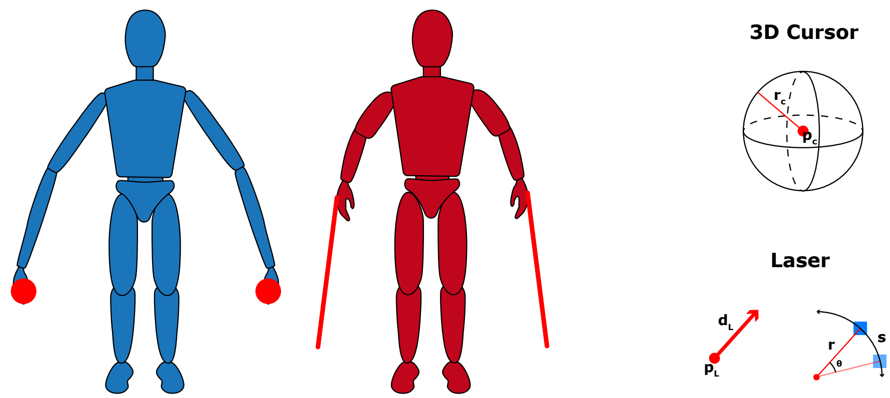
We ask the following research questions with our study:

- **RQ1**: Can users learn to control two bodies to accomplish a simple handoff task between them?
- **RQ2**: How does the position of the third person body affect user's performance in this task?
- **RQ3**: What strategies do users employ while solving this task?
- **RQ4**: What is the embodiment relationship that users have with each body?
- **RQ5**: Do users have awareness of their arms that are not involved in the handoff task?

4.1 Study Design

Our study is designed such that it requires the user to manipulate two virtual bodies to complete a handoff. We vary cube size, distance and orientation between the two bodies, as well as which body type is controlled in the first versus third person. We selected a simple task that could be performed repeatedly in different configurations to understand the impact of our independent variables on task performance.

4.1.1 Bodies. The user is given two bodies with differing morphologies and functionalities. Interestingly, such a contrast splits



(a) Left: The Stretch Body has arms 2 times the length of the user's arms. Interaction happens with 3D cursor at the hands. Right: The Laser Body has arms that match the user's. Interaction happens with laser interactors attached to the hands.

(b) Geometries associated with 3D cursor and ray based interaction used by the two bodies.

Figure 2: Diagrams of the bodies used in our experiment, and the geometries associated with their interactions

the affordances in virtual space to each hand while still being a bimanual coordination task from the user's control. The **Stretch Body** uses a 3D cursor attached to each hand, so it is considered a positioning interaction. The body's arms are scaled to be two times the length of the user's, giving the user an extended range for grasping similar to Go-Go [44]. This is done to make the handoff task easier, increasing the space where the user can perform the handoff (see Figure 3). This body is rendered in blue. The **Laser Body** uses ray-based interaction from each hand, pointing in the direction the hand is pointed. Ray-based interaction can be performed from a range of distances, allowing us to vary this distance as a parameter of the study design. This body is rendered in red. Diagrams of these bodies can be seen in Figure 2.

While other selection technique choices for the two bodies are possible, these were chosen for pragmatic reasons. We considered using two bodies that both utilize positioning interactions (such as two Stretch Bodies) as this would more closely resemble how two humans would hand off an object between each other. In order to vary the position of one of the bodies while also allowing their range of motion to still intersect (see Figure 3), we would need to vary the length of the arms of one of the bodies. This would introduce a new gain applied to the positioning interface for certain positions of the body, making the study more difficult for participants to learn. A pointing interaction was chosen for the second body due to gains remaining constant at different distances. Our results shed light on how task performance is shaped by a heterogeneous control design that assigns asymmetrical affordances through a split in first and third person perspective, differing dual body morphology (stretch and laser body), and functionality (reach and aim).

4.1.2 Scenario. The VE has a pedestal with a cube upon it. Opposing the pedestal is the Stretch Body. The user's task is to use the Laser Body to pick up the cube and hand it off to the Stretch Body. The layout diagram of this scenario is depicted in Figure 4. This scenario is performed twice, once with the Stretch Body as the first person body and the Laser Body as the third person body, and then again with the Laser Body as the first person body and the Stretch Body as the third person.

There is an axis between the Stretch Body and the pedestal, which we refer to as the central axis. The length of this axis, d on the layout diagram, is 2m. To facilitate a handoff between the two bodies, the Laser Body is placed halfway along the central axis, and offset perpendicular to it. This placement is to help facilitate handoff during the experiment (See Figure 3). The Laser Body faces the midpoint of the central axis by default.

The user is asked to perform 40 handoff trials. Each trial has a different set of conditions. The independent variables for these conditions are as follows.

Cube Size (Size). We predict that target size will have an impact on task completion time. To confirm this we use two target object sizes.

Laser Offset (Offset). The mapping of rotational hand motions to translational object motions is a function of the distance along the laser that the object is grabbed. To investigate how this might effect performance, the Laser body changes its distance from the central axis at distances $\pm 1.2m$, $\pm 2.4m$ (4 total). These values are chosen so that the third person body does not appear far out of the user's point of view, reducing the need for users to scan the scene at the start of each trial. The laser offset variable provides a way for determining the overall task difficulty of both phases (laser portion, handoff portion).

Third Person Body Rotation (Rotation). We apply a rotation to the third person body in each trial. The body starts by facing the central axis and is rotated relative to this default orientation by 0, ±22.5, ±45 degrees (5 total). These values represent an even sampling of a 90 degree, forward facing arc. Pilot studies showed that rotations beyond this 90 degree value create situations where the ergonomics of the task become difficult to navigate.

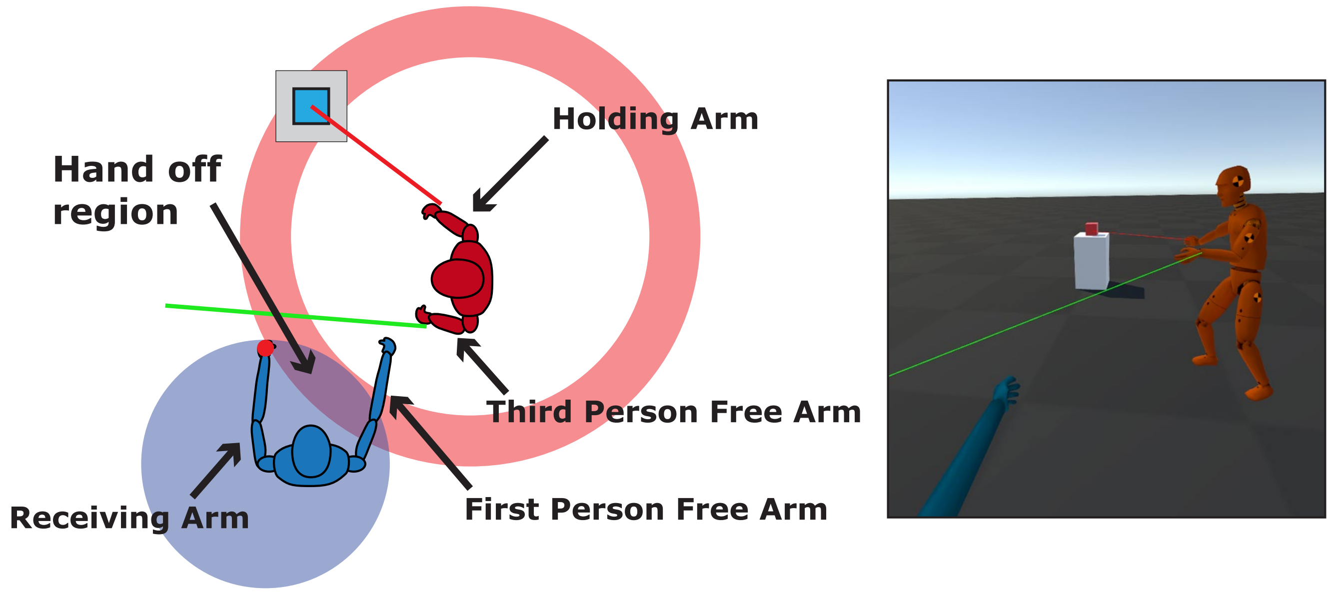


Figure 3: Diagram showing the reason for the specific geometry of the Laser Body placements. When the Laser Body grabs an object, there is effectively a ring around the body that the the object can be moved to. This ring intentionally intersects a region of space near the Stretch Body, called the "handoff region", so that the user can control the Stretch Body to grab the object. On the right is a screen capture of this scenario as seen by participants in the study.

During pilots of our study, we tried a number of values for these specific parameters. We settled on the numbers presented for providing a balance between ease and difficulty, and also needing to limit the number of conditions such that the experiment could be completed in under one hour.

We create random permutations of each of these variables to form the conditions of each trial for each user. No user saw the trials in the same order.

When the task begins, the user picks the cube off the pedestal using the Laser Body. We refer to this as the **laser portion** of the task. Then the user must get the cube into either of the Stretch Body's hands as quickly as possible, which we refer to as the **handoff portion**. We define the following terms to help facilitate discussion about what each hand is doing during this process.

Holding Arm This is the arm on the Laser Body that was used to pick up the target object off the pedestal.

Receiving Arm This is the arm on the Stretch Body that was used to grab the object from the laser.

Third Person Free Arm This is the arm on the third person body that is not involved in the handoff.

First Person Free Arm This is the arm on the first person body that is not involved in the handoff.

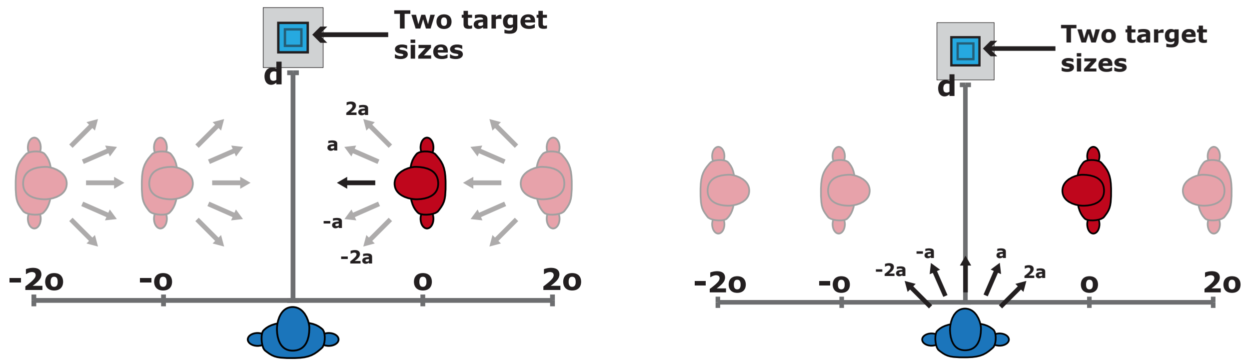
Upon handoff completion, a sound is played and a UI element appears instructing the user on how to start the next trial. In the Stretch Scenario users controlled the Stretch Body from the first person and the Laser Body from the third person. In this scenario, the Holding Arm is a third person arm, and the Receiving Arm is a first person arm. As the Laser Body is the third person body, it was rotated by the Rotation variable. In the Laser Scenario users controlled the Laser Body from the first person and the Stretch Body from the third person. The Holding Arm was a first person arm, and the Receiving Arm was a third person arm. The Stretch Body was rotated by the Rotation variable. The function of the bodies remained constant in both scenarios so that we could more easily compare results between the scenarios. If we had changed the length of the Stretch Body's arms between scenarios, for example, users would not only have to learn to control the body from a new perspective, but also learn a new control gain.

4.2 Apparatus

The system used to run this study was a desktop computer running Windows 10, connected to an Oculus Quest 2 VR headset using an OculusLink cable. We provide an approximate body reconstruction from tracked controllers and VR headset using inverse kinematics (IK). Our IK solution is the VRIK Unity Package, which uses the RootMotion FinalIK [46] middleware to compute inverse kinematics. The study was done standing in a 4 by 4 meter space.

4.3 Participants

We recruited nineteen participants from university mailings lists, 6 women and 12 men, 1 chose to not respond to this question. Participants included undergraduate students, graduate students, and campus staff. The mean age was 22 years old, ranging from 18 to 32 years. Sixteen participants were right handed, 3 left handed. 7 participants had never used VR at all, 10 participants had used VR a



(a) Stretch Scenario Setup. In this scenario, the Stretch Body is controlled from the first person, and the Laser Body from the third person.

(b) Laser Scenario Setup. In this scenario, the Laser Body is controlled from the first person and the Stretch Body from the third person.

Figure 4: Diagrams of both experimental setups. Distance from the Stretch Body to the target is d. Laser Body is offset by several factors of o from the center line. For the Stretch scenario, the laser bodies rotate $\pm a$, $2a$ from it's center direction. For the Laser scenario, the Stretch Body has the same rotational offsets. Each scenario has two target object sizes. This provides a total of 80 conditions tested.

few times, 1 participant had used VR many times, and 1 participant uses VR on a regular basis.

4.4 Study Procedure

Participants completed a background and demographic questionnaire. Each participant received the same training on how to use a VR headset, regardless of their experience with VR. Participants then entered a training scenario designed to teach them the basics of interaction using our software. A calibration process for the IK software was also run at this time. Then participants were taught to grab objects from the first person body, the third person body, and then taught how to hand objects off between them. Participants were allowed to spend as much time in this scene as they wanted, until they self reported feeling comfortable with how interaction worked.

We then loaded participants into one of the two experimental scenarios and allowed them to practice the study task as many times as they wanted. Participants would self report once they felt comfortable to be able to complete the task. After their practice, we reloaded the scene and began the experiment. Users were instructed to perform the handoffs as quickly as they could. We always insert 1 handoff task at the beginning of the condition list which we did not collect data for. The completion of this first condition would allow the participants to control when the experiment began. After each experiment, participants were removed from the virtual environment and filled out a questionnaire asking about embodiment as well as qualitative questions about their experience. Study facilitators also used this time to ask them specific questions about the experiment based on observations.

Participants performed this entire procedure twice, once for each scenario. Scenarios were presented to users in random order. After the last scenario they answered additional qualitative questions about the whole procedure. They were compensated with a $25 gift card.

4.5 Measures

We record **task completion time** as the dependent variable during handoff tasks. Performance times are chosen so that we can compare across different experimental setups to find how relative body positioning affects user's ability to complete tasks. Instructing participants to perform tasks as quickly as possible places an emphasis on efficiency. Because users are operating on efficiency, this allows us to investigate what strategies they use to complete the task and how those might evolve over time. Tracking this development of strategy tells us something about the learning curve for the interaction, and whether or not efficient strategies are discoverable. Task time begins as soon as a trial starts and ends as soon as the handoff is complete.

To understand what the hands are doing during this study, we record position and rotation data from the VR controllers at each timestep of the experiment. This data is processed post-hoc to compute discrete controller velocity over time data for each trial. We use this hand tracking data to classify the strategies that users used during the study (see Figure 7).

Embodiment questions were taken from Peck and Gonzalez-Franco's standardized questionnaire for avatar embodiment [42]. Questions Q1-Q3, Q6-Q9, Q15-Q18, and Q20 were asked. We selected questions from the categories Body Ownership, Agency and motor control, Location of the body, and External appearance, as those were the categories we were most interested.

We noticed a trend during our pilot studies about people's awareness of their First Person Free Arms, so after each participant completed the Stretch Scenario they were asked the following question:

"Imagine a trial where you reached out with the right hand of the Laser Body, grabbed the cube, brought it in front of your Stretch Body. You reach out with the left hand of the Stretch Body and take the cube. What is the right hand of the Stretch Body doing?"

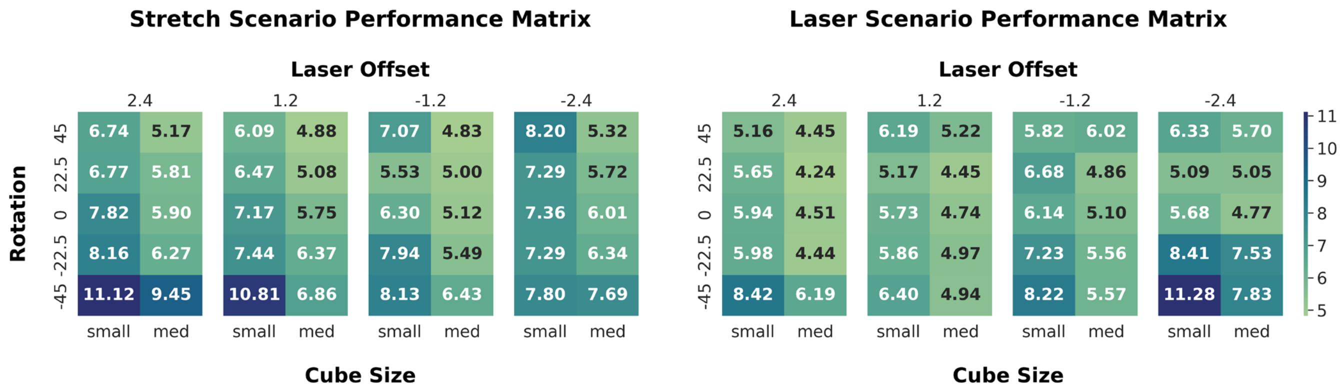


Figure 5: Heat map representation of task completion times. Each individual cell depicts mean task completion time across users for each combination of independent variables.

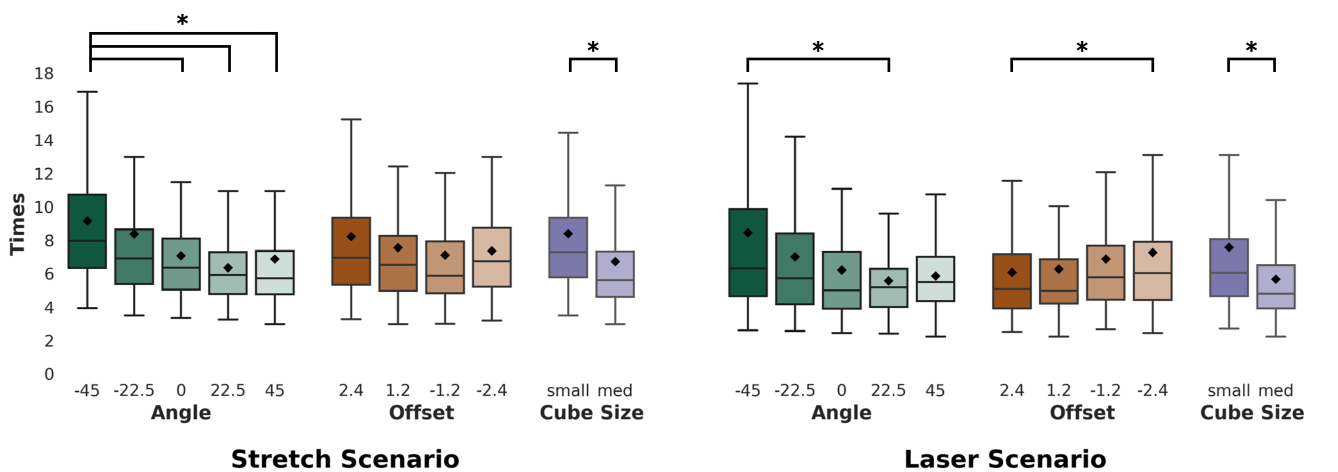


Figure 6: Completion time across all trials, split by independent variable and scenario. Means are represented by the diamonds.

We refer to this question as the **arm awareness question**. Answers to this question were coded by study facilitators for post-hoc analysis.

5 RESULTS

5.1 Stretch Scenario Quantitative Results

It took participants between 2.98 and 85.89 seconds to complete each task across all conditions, with a mean time of 7.56 seconds and a standard deviation of 5.17 seconds. A repeated measures ANOVA was performed on completion times on the within subject factors: Size, Offset, and Rotation. Mauchly's test showed that the assumption of sphericity had not been violated. Mean task performance times are presented for each condition in Figure 5 (left), and task performance time split by variable is presented in Figure 6 (left).

Rotation × Offset A two way interaction was found between Rotation × Offset ($p < 0.05$). Paired t tests revealed that Rotation values had a significant effect on completion times (Figure 10b). After applying Bonferroni corrections ($\alpha = 0.0025$), significant differences were found when the Laser Body was rotated toward the Stretch Body at three of four Offset values. For an offset value of 2.4, there was a significant difference between Rotations of -45 and 0 ($p < 0.001$), -45 and 22.5 ($p < 0.001$), -45 and 45 ($p < 0.001$). For an Offset of 1.2 there was a significant difference between Rotations of -45 and 22.5 ($p < 0.001$). For an Offset value of -1.2, there was a significant difference between Rotations of -45 and 22.5 ($p < 0.001$).

To summarize, completion times were generally faster across all offsets for positive values of Rotation, but this effect is lessened for positive Offset values.

Main Effects A main effect of Size was found ($p < 0.001$) with completion times being significantly higher for small target sizes. This difference is expected for target acquisition tasks and is explored in the discussion.

A main effect was found for Rotation ($p < 0.001$) with completion times being significantly higher for negative values of Rotation. After applying Bonferroni correction ($\alpha = 0.01$), significant differences were found between Rotation values of -45 and 0 ($p < 0.001$), -45 and 22.5 ($p < 0.001$), -45 and 45 ($p < 0.001$), -22.5 and 22.5 (p <

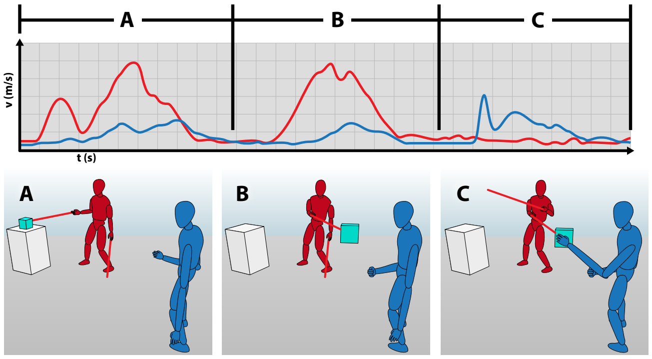


Figure 7: Example of how a velocity plot can be used to reconstruct the strategy used by a user. This example depicts the most common strategy, Grab From Laser. Top: A velocity over time plot used to classify each strategy. Red lines indicate velocity over time of the Holding Arm, and blue lines indicate velocity of the Receiving Arm. Section A represents the laser portion of the task, and sections B and C represent the handoff portion. The plot (and task) end as soon as the user grabs the cube with the Stretch Body. Bottom: A graphical depiction of the bodies performing the strategy. A: User grabs the cube with the right hand of the Laser Body. B: User moves the cube in front of the Stretch Body. C: User reaches out with the left arm of the Stretch Body and completes the handoff.

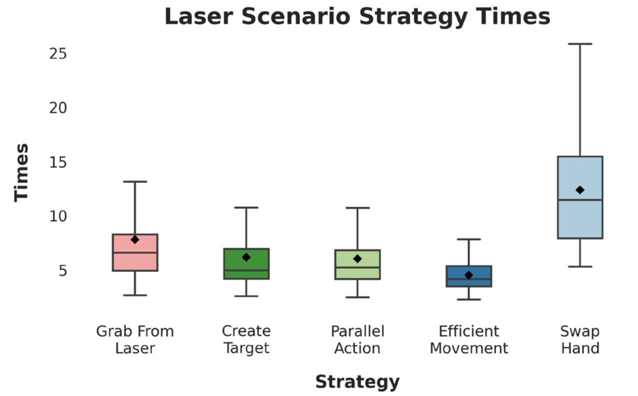


Figure 8: Performance times of each strategy used during the Laser Scenario. Means are represented by the diamonds.

	Usage					
Strategy	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Mean Time**
Grab From Laser	82	75	65	35	31	7.8 ± 5.4s
Create Target	3	8	24	38	48	6.2 ± 3.5s
Parallel Action	40	35	30	37	28	6.0 ± 3.2s
Efficient Movement	18	26	26	41	44	4.5 ± 1.7s
Swap Hand	9	8	7	1	1	12.4 ± 5.5s

Table 1: Quantitative results of strategy selection. Usages are split and counted in 5 quintiles. Mean time and standard deviation was computed across all uses of each strategy.

0.01). These statistics suggest that negative values of Rotation have a significant effect on completion times.

5.2 Laser Scenario Quantitative Results

It took participants between 2.24 and 56.58 seconds to complete each task across all conditions, with a mean time of 6.63 seconds and a standard deviation of 4.4 seconds. A repeated measures ANOVA was performed on completion times on the within subject factors: Size, Offset, and Rotation. Mauchly's test showed that the assumption of sphericity had not been violated. Mean task performance times are presented for each condition in Figure 5 (right), and task performance time split by variable is presented in Figure 6 (right).

Angle × Cube A two way interaction was found between Rotation × Cube ($p < 0.05$). Paired t tests revealed that Cube and Rotation values had a significant effect on task completion times (Figure 10a). After applying Bonferroni corrections ($\alpha = 0.005$), significant differences were found between all values of Cube for each value of Rotation : -45 ($p < 0.005$), -22.5 ($p < 0.005$), 0 ($p < 0.001$), 22.5 ($p < 0.001$), 45 ($p < 0.001$). In summary, task completion times were significantly faster for the larger Size target.

After applying Bonferroni corrections ($\alpha = 0.005$), significant values were also found for Rotation when Size is held constant. For the small target, significant differences were found between Rotation values of -45 and 0 ($p < 0.005$), -45 and 22.5 ($p < 0.001$), -45 and 45 ($p < 0.001$). For the medium size target, significant differences were found between Rotation values of -45 and 0 ($p < 0.01$), -45 and 22.5 ($p < 0.001$), -45 and 45 ($p < 0.001$).

To summarize, completion times were generally faster across all Sizes for positive values of Rotation, and this effect has a more significant impact if the Size is small.

Main Effects. A main effect of Size was found ($p < 0.001$) with completion times being significantly higher for small target sizes. This difference is expected for target acquisition tasks and is explored in the discussion.

A main effect was found for Rotation ($p < 0.001$) with completion times being significantly higher for negative values of Rotation. After applying Bonferroni correction ($\alpha = 0.01$), significant differences were found between Rotation values of -45 and 0 ($p < 0.001$), -45 and 22.5 ($p < 0.001$), -45 and 45 ($p < 0.001$), -22.5 and 22.5 ($p < 0.001$), -22.5 and 45 ($p < 0.01$). These statistics suggest that negative values of Rotation have a significant effect on completion times for the Laser Scenario.

5.3 Stretch Strategies

We looked at velocity over time plots of participant's hands to determine which strategy they employed during each scenario. The strategies manifest on these plots very clearly because each step of the strategy often involves the ordering of which hand is moving when. An example of how this classification is performed can be seen in Figure 7.

During the Stretch Scenario, users primarily employed a single strategy, which we call **Grab From Laser**.

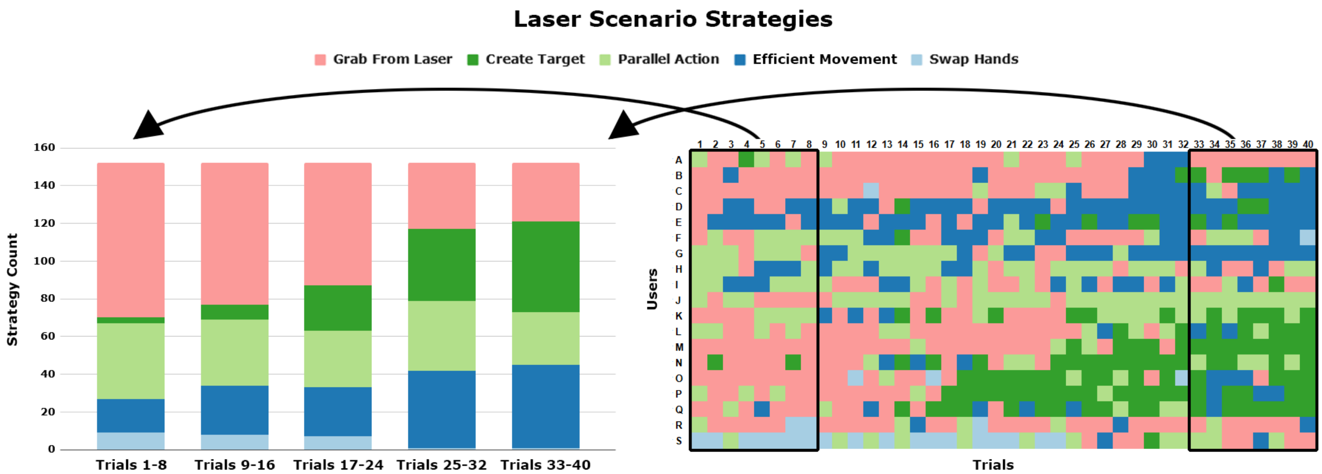


Figure 9: Strategies employed during the Laser Scenario portion of the user study. Left: Strategies put into 5 bins showing overall strategy development over time. Right: Matrix plot of all users and trials. Each row is an individual user. Each column is a trial presented in the order that the user saw them. Presented this way, the x axis represents time across a single run of the Laser Scenario.

- Pick up cube with the Holding Arm
- Bring cube in front of the Stretch Body
- Grab cube with Receiving Arm

Some users would preempt the placement of the cube in front of the body by beginning to reach their hand forward as the cube approached. Some users waited until the cube was stopped to reach out for it.

5.4 Laser Strategies

Users employed 5 unique strategies in the Laser Scenario. To better understand how users developed their strategy use over time, we split the trials into quintiles (5 blocks with 8 trials per block). This number of blocks was chosen so we could observe strategic development at the start, middle, and end of the trials, while having one data point in between each to help interpolate the strategic development. Unique instances of each strategy were counted in each quintile and are presented in Figure 9 (left) and Table 1. We also present the raw data for each user in Figure 9 (right) and which strategy they employed on which trial. Performance times for these strategies can be found in Figure 8 and Table 1. Welch's t-test was performed pairwise against all strategies, and significance was found ($p < 0.001$) for every pair except for Create Target and Parallel Action, where no significance was found.

The same Grab From Laser strategy was present in this scenario. We present the following 4 additional strategies that emerged.

Create Target

- Pick up cube with the Holding Arm
- Move the Receiving Arm outward away from their body
- Bring the cube into the proximity of the Receiving Arm
- Grab with the Receiving Arm. Sometimes minor adjustments of Holding and Receiving Arms are required

Parallel Action

- Pick up cube with the Holding Arm
- Simultaneously reach out the Receiving Arm, and bring the cube into the proximity of the Receiving Arm
- Grab with the Receiving Arm. Sometimes minor adjustments of Holding and Receiving Arms are required

Efficient Movement

- Pick up cube with the Holding Arm
- Bring the cube into the proximity of the Receiving Arm
- Grab with the Receiving Arm. Sometimes minor adjustments of Holding and Receiving Arms are required

Swap Hands

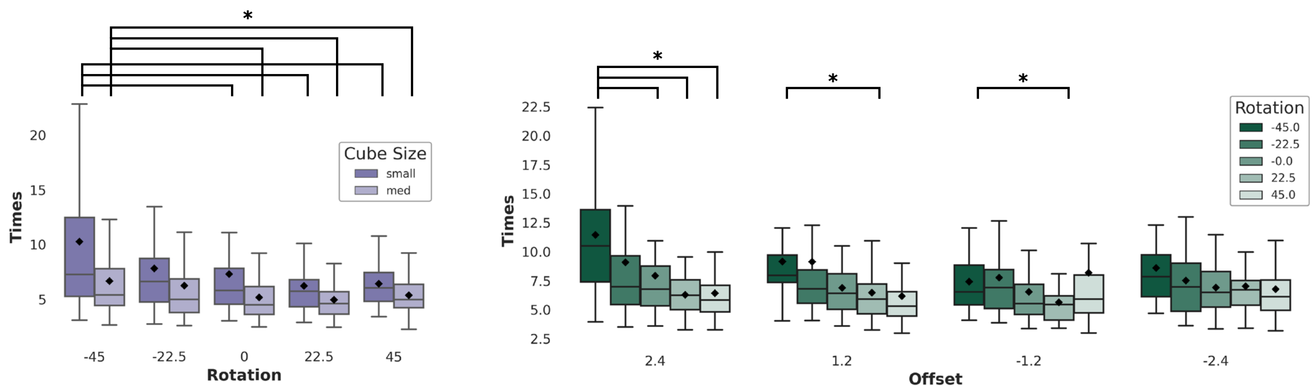
- Pick up cube with the Holding Arm
- Swap which hand is the Holding Arm
- Bring cube in front of the Stretch Body
- Grab cube with Receiving Arm

This is the least common strategy that was primarily employed by one user. Ultimately this user changed their strategy in the study, remarking "this doesn't seem very efficient".

5.5 Embodiment Ratings

Each of our embodiment questions were asked about the user's first and third person bodies, and each user answered each set of questions twice, one for each scenario. Responses to the question were collected on a 7 point Likert scale, and ranged from -3 to 3. Positive scores in this analysis represent agreement with the question, and negative scores represent disagreement.

We computed scores for different categories of embodiment from [42]. Numbers reported are means across all users. For the first person body, users reported a Body Ownership score of 2.47, and Agency and motor control score of 5.52, a Location of the body score of 0.02, and an External appearance score of -2.65. For the third person body, users reported a Body Ownership score of -1.34,



(a) (b)

Figure 10: (a) Plot showing interaction between Cube × Angle for the Laser Scenario. (b) Plot showing interaction between Offset × Angle for the Stretch Scenario. Means are represented by the diamonds.

and Agency and motor control score of 3.58, a Location of the body score of -0.45, and an External appearance score of -2.95.

Using a modified analysis of computing an a total score from [42], users reported an overall embodiment score for the first person bodies of 0.35, and an overall embodiment of the third person bodies of 0.02.

The raw data of user answers can be found in Figure 11.

5.6 Arm Awareness

When we asked users the arm awareness question, 10 users responded that they were not sure what the arm was doing. 5 users issued an incorrect answer, each of which assumed that the arm was hanging stationary at their side. 4 users issued correct answers, that the arm was following the motions of the Laser Body's right arm. 3 of these users had seen the Stretch Scenario **after** the Laser Scenario, and when asked to explain how they knew what the arm was doing, they explained they had seen the Stretch Body's right arm following their own during the previous scenario, and that they were inferring that it must be doing the same thing during this scenario even though they were not actively aware of it. The last user that answered correctly explained that they were actively aware of what the arm was doing during their scenario.

6 DISCUSSION

6.1 Quantitative Results

Our research questions asked whether or not people can complete this task (RQ1) and what are the main factors affecting performance in this task (RQ2). As all users were able to accomplish each task, the answer to RQ1 is that users are able to learn to control two bodies to complete handoff tasks. The **cube size** and **third person body rotation** independent variables had a significant ($p < 0.001$) impact on performance times.

6.1.1 Size. It is not surprising that target size has an impact on performance. This task resembles a multi-part Fitts' Law [12] task, in particular the laser portion of the task could be modeled directly by Fitts' Law, as it is a pointing task, suggesting that target size would indeed be a factor in performance. The handoff portion is more complicated because the user is simultaneously moving the target with one hand while trying to select it with the other. We leave a deeper investigation of the correct modeling of the handoff task to future work.

6.1.2 Reference Frames. We explored two components of modifying the reference frame between the two bodies: moving the Laser Body relative to the Stretch Body (Offset variable) and rotating the third person body relative to the first person body (Rotation variable).

Statistical analysis of the Offset variable did not show to have a significant effect on task completion times. This implies that the values chosen do not have significant impact on pointing performance in the first or third person.

Statistical analysis of the Rotation variable provides a compelling answer for RQ2, and showed that it is likely the most significant factor on performance in this task, with rotations towards the first person body (negative values of the variable) being significantly more difficult than rotations away. This is likely due to the fact that as the third person body gets rotated toward the first person body, the movement axis for the user's hands gets flipped. For example, physical motions of the hand to the left, when synchronized onto a body that is facing you, will move it's hands to the right (relative to the first person view). Although these results don't identify the optimal placement of bodies for this task, they do suggest that an optimal placement could exist, and can be used to inform future work on investigating this placement.

6.2 Strategy Choices

We ask what types of strategies users employ while completing tasks in this study (RQ3). Users primarily used one strategy during the Stretch Scenario. We looked at whether or not users started with different strategies depending on the order that they completed the scenarios and found no correlation.

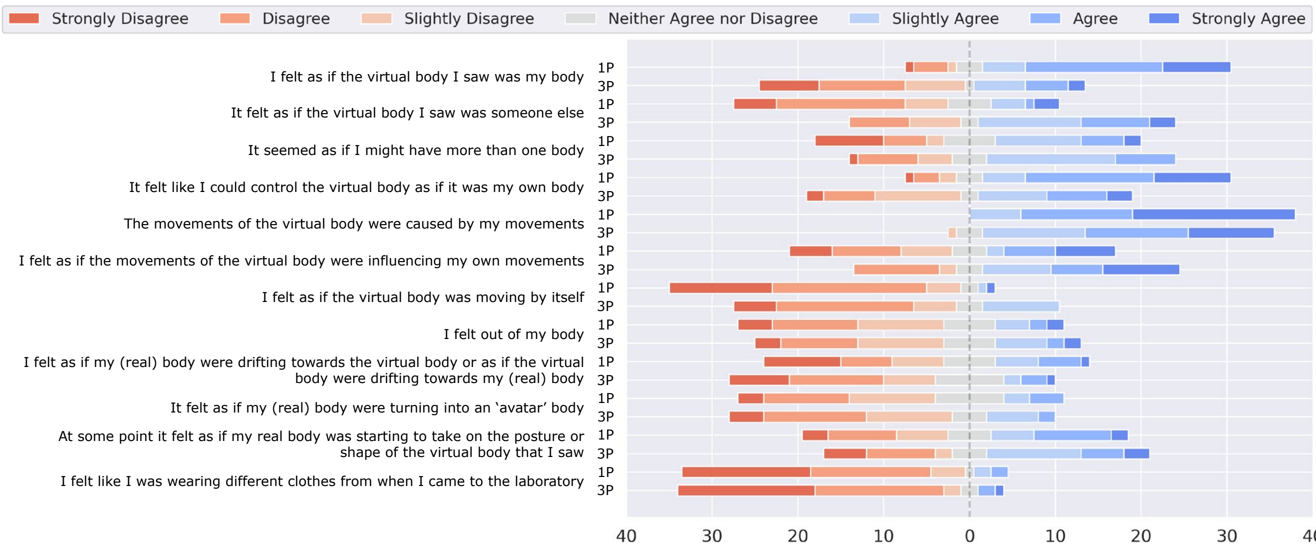


Figure 11: Answers to our embodiment questions. Each question was asked about the different first and third person bodies that the user controlled.

The Laser Scenario provides an interesting answer to RQ3. As seen in Figure 9, participants tended to start with the Grab From Laser strategy. We believe this is because this strategy most closely resembles the prompt given to the participants: Pick up the cube with the Laser Body and hand it off to the Stretch Body. Many users also started with Parallel Action, a very similar strategy. These two strategies account for 80% of strategies used in the first quintile of trials.

We observe that over time, the Grab From Laser strategy becomes less popular with participants. Observations from study facilitators may explain this. We noticed that users had a more difficult time controlling the Receiving Arm in the third person than controlling the Holding Arm in the third person. In the study we observed two behaviors for learning the mapping transforms of these bodies. In the Stretch Scenario, users would learn the mapping transform of the Laser Body by swinging their Holding Arm around to determine which way they should be moving their physical arm to point at the target. In the Laser Scenario, users tried a similar action to learn the mapping transform of the Receiving Arm of the Stretch Body, but struggled quite a bit with actually targeting something after doing this process. We believe this may be because **translational motions are more difficult to map into different rotational reference frames**, and that this difficulty may contribute to their avoidance of Grab From Laser.

Another observation is that once users moved away from Grab From Laser, they would often stick to whatever new strategy they encountered. These new strategies were usually Create Target or Efficient Movement. 60% of users ended the study using one of these two strategies, as opposed to 14% of users starting with it. These were also the two highest performing strategies (see Figure 8).

6.3 Embodiment vs Extended Cognition

To answer the question of what the embodiment relationship is to each body (RQ4), we look at the self reported embodiment scores. The scores indicate that participants did embody the first person body, and did not embody the third person body. For the first person body, users on average reported high Body Ownership, Agency, and Motor Control scores. For the third person body, users reported low Body Ownership, high Agency, and middle Motor control scores. Low body ownership scores can be attributed to the body being in a different reference frame, and the presence of a first person body in both scenarios. High agency scores implies that the users felt like their actions were producing the motions on the third person body. This result agrees with the result in Gorisse et al. [17], where they also found a high sense of agency of third person avatars. Kondo and Sugimoto [27] also explore the concept of body ownership as it applies to controling multiple synchronized bodies. Their results show that body schema in users does not change when controlling multiple bodies, but does change when controlling a split body. The motor control scores can be explained by some of the qualitative feedback we received. We asked an open ended question about which scenario they found easier to perform, and 12 out of 19 users indicated that they found the third person bodies more difficult to control. Future studies could further explore this by having strict experimental controls on difficulty between first and third person bodies, given that a perceived asymmetrical difficulty level of control could account for users not reporting a sense of embodiment despite reporting a sense of agency.

Based on these results we can confidently say that users do not embody the third person body, but the question remains as to what their relationship is to this body, and how they are able to use it effectively to accomplish handoff tasks. Two features of Embodied Cognition that are most relevant here are the idea that our sensory experiences are *embodied* within the interactions of our bodies and the physical environment [15], and that cognition is *extended* beyond an intracranial locus which also encompasses the coupling between the tools we use [1]. Although our users do not report a sense of embodiment, the reported higher levels of agency and some level of motor control is indicative of their cognition being extended through their interactions with the tool and design we propose. Someone's cognition being embodied is not solely dependent on a user reporting an active awareness or perceived sense of embodiment. The relationship between a user's intended goals, their behaviors, and the coupled interactions in their environment (including a virtual space [47]) with the performance scores they produced is what demonstrates [47] an embodied cognition.

Dourish writes in Where the Action Is [10] about how Phenomenology can explain the relationship between how we operate a mouse and what we do with a mouse. Such foundational work in HCI is shown to have already been making the connections of what eco-psychological components constitutes the entirety of a user interaction. Given that a user frequently interacts through the mouse, the mouse becomes an extension of the hand as we perform actions in screen based interfaces. Therefore, the mouse is said to be *ready-at-hand* in terms of Heidegger's phenomenology. Contemporary works in Embodied Cognition have elaborated on that notion, including empirical studies that break down spatio-temporal variables of a user acting through and or with the tools embedded in their environments for which a person's tool use enacts their intended goals [9]. Bergstrom et al. [3] have shown that this concept of tool extension can be measured in virtual reality systems, and applying their methodology to measure the level of tool extension in our experimental setup could be a promising next step to verify this.

These findings show that users are controlling the third person bodies in a similar way to how we use tools, by extending our bodies through the tools to perform the given task. Users appear to be able to control virtual tools with body-like affordances from many different positions, and can use these body-tools to perform tasks requiring coordination.

6.4 Body Awareness Questions

We are interested in whether or not users have awareness of their unused arms not involved in the handoff task (RQ5). The answer to this question is that most users were not aware of what the First Person Free Arm was doing during the Stretch Scenario. Some users even reported incorrectly that the arm was doing nothing, hanging at their side. We have included an illustration from a user's point of view performing the task as described in Figure 12. The arm is clearly visible and is following the motions of their physical arm, yet most users were not actively aware of it. While we are not able to provide an explanation for this, we believe two potential concepts from cognitive science may point us in a direction: selective attention and anticipation.

Theories on Selective Attention can help inform this mismatch of user reports and the actuality of their behaviors. Selective attention is the explanation for how our perceptual systems ignore some stimuli and focus on others depending on the task that we are performing. Research exists showing this phenomenon can occur

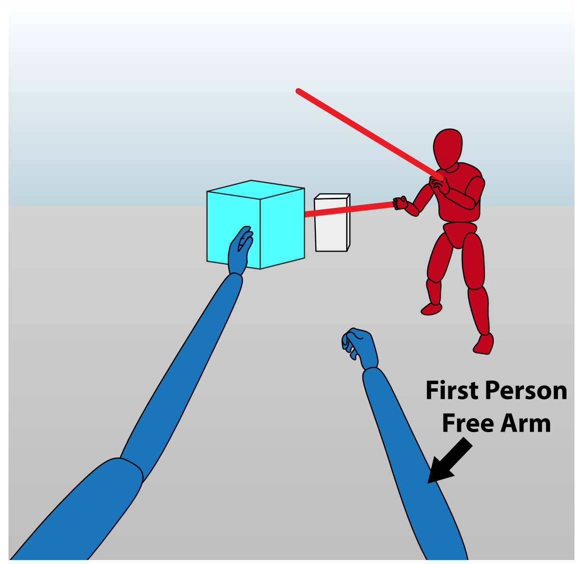


Figure 12: Illustration of the first person view when asked the Arm Awareness Question. When users were asked what their First Person Fee Arm (highlighted) was doing during a handoff task, most reported that they did not know, despite it being in clear view.

during video watching [51] and in VR [35]. Alfred Yarbus [58] also explores this concept in Eye Movements and Vision. He shows an example people's tracked eye positions while looking at a photograph. The eye's trajectory changes based on the objects in the photograph they are instructed to analyze. There is some aspect of our sensory system that only perceives objects relevant to our current task. In The Humane Interface [45], Jef Raskin writes about our *locus of attention*, which is the object or idea that you are currently actively thinking about. Raskin identifies this concept as the thing that prevents us from seeing things that are in plain view.

We can see from two of the strategies employed during the Laser Scenario (Create Target, Parallel Action) that users take preemptive action with their receiving hand, as if they are anticipating the coordination that must occur. Modeling behind the physical and cognitive mechanics of bimanual coordination have brought forth such a notion of anticipation, to explain behaviors that do not require active awareness of subcomponent mechanics for which an observer can sometimes assume as an inferential process [54]. That is to say that users do not need to create mental models predicting and updating where their hands will go, but rather the interaction with their virtual environments continually defines their ecological affordances which emerges as the observed trajectory of user coordination [7, 13].

Franz Mechsner highlights related examples in sports science when looking at the performance and coordination of athletes [37]. He argues that spontaneous bimanual interaction may happen at a perceptual-cognitive level, and therefore people do not need to plan their motor commands as an additional process. The implication being that motor control and planning can happen at a level in our brains outside of active awareness.

If the user is not actively aware of the First Person Free Arm's location, then where do users perceive their arm? Our study does not directly answer the question, but we hypothesize that it may be located in the position of the Holding Arm. Users appear to be able to control which arms of which body they act through and attend to in a fluid manner, deciding to control a given arm on a given body as the tasks sees fit. We hypothesize that it is the task that determines how users perceive the arms.

7 LIMITATIONS

Our user study focuses on the exploration of experimental parameters that affect the user's ability to coordinate with multiple bodies. This is set up with both independent variables of body perspective and morphology. Future experiments should test user embodiment using more than questionnaire results, such as using physiological measures like eye tracking, or task driven behavioral measurements such as in the rubber hand experiment [5].

All of our results should be interpreted in the context of the affordances we provide users (stretched arms, ray based interaction). There are two possible limitations of our study as they relate to body morphology. Our arm awareness results could be due to the fact that the user was focused on controlling a body that functions very different than their own. A future study could look at this by providing a similar task between two bodies with a regular homogenous morphology. Additionally, the results could also be due to the specific parameterization chosen for the bodies (stretched arms). Further explorations of this phenomena should also be examined under different parameterizations of the bodies provided, such as longer and shorter arms.

8 CONCLUSION AND FUTURE WORK

In this work we have examined a study where participants controlled two bodies in VR to perform a coordinated task. Users controlled a laser body to pick up a cube and hand it off to a body with long arms. This task performed 40 times with varying positional and rotation offsets to measure the ideal body locations for such a task. The whole procedure was performed twice, once from the view of each body. Our findings show that rotation is the most significant factor relating to performance. We also provide qualitative results that suggest that users do not embody the third person body during the study, but that they do possess some ownership and agency over it. Such results connected to concepts of embodiment in Cognitive Science to elaborate in how such phenomena could be understood. We also find that user attention during the task may change their active awareness of body location.

This is only an initial foray into studying coordination across multiple bodies in VR. We hope this will serve as an impetus to further study this phenomenon. Further studies should continue to develop the connections to anticipation and extended cognition. From an interaction perspective, now that we have evidence for coordinated bimanual action across bodies with different body schema, future work should look into how this phenomenon can be harnessed for more productive interaction techniques.

ACKNOWLEDGMENTS

This work was in part supported by a gift from Accenture. We thank them for their support, and we thank our research colleagues in Accenture Labs for sharing their expertise and feedback throughout the process.

REFERENCES

[1] Ken Aizawa. 2014. Extended cognition. In *The Routledge handbook of embodied cognition*. Routledge, 49–56.
[2] Ravin Balakrishnan and Ken Hinckley. 2000. Symmetric Bimanual Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands) *(CHI '00)*. Association for Computing Machinery, New York, NY, USA, 33–40. https://doi.org/10.1145/332040.332404
[3] Joanna Bergström, Aske Mottelson, Andreea Muresan, and Kasper Hornbæk. 2019. Tool Extension in Human-Computer Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3290605.3300798
[4] Frank Biocca. 1997. The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments [1]. *Journal of Computer-Mediated Communication* 3, 2 (09 1997). https://doi.org/10.1111/j.1083-6101.1997.tb00070.x JCMC324.
[5] Matthew Botvinick and Jonathan Cohen. 1998. Rubber hands 'feel' touch that eyes see. *Nature (London)* 391, 6669 (1998), 756–756.
[6] W. Buxton and B. Myers. 1986. A Study in Two-Handed Input. *SIGCHI Bull.* 17, 4 (apr 1986), 321–326. https://doi.org/10.1145/22339.22390
[7] Umberto Castiello. 1999. Mechanisms of selection for the control of hand action. *Trends in Cognitive Sciences* 3, 7 (1999), 264–271. https://doi.org/10.1016/S1364-6613(99)01346-7
[8] Mark R Costa, Sung Yeun Kim, and Frank Biocca. 2013. Embodiment and embodied cognition. In *Virtual Augmented and Mixed Reality. Designing and Developing Augmented and Virtual Environments: 5th International Conference, VAMR 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part I 5*. Springer, 333–342.
[9] Rick Dale, Riccardo Fusaroli, Nicholas D Duran, and Daniel C Richardson. 2013. The self-organization of human interaction. In *Psychology of learning and motivation*. Vol. 59. Elsevier, 43–95.
[10] Paul Dourish. 2001. *Where the Action is: The Foundations of Embodied Interaction*. MIT Press, Cambridge, MA, USA.
[11] Ylva Fernaeus, Jakob Tholander, and Martin Jonsson. 2008. Towards a New Set of Ideals: Consequences of the Practice Turn in Tangible Interaction. In *Proceedings of the 2nd International Conference on Tangible and Embedded Interaction* (Bonn, Germany) *(TEI '08)*. Association for Computing Machinery, New York, NY, USA, 223–230. https://doi.org/10.1145/1347390.1347441
[12] Paul M. Fitts. 1954. The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *Journal of Experimental Psychology* 47, 6 (1954), 381. https://doi.org/10.1037/h0055392
[13] Jonathan B Freeman, Rick Dale, and Thomas A Farmer. 2011. Hand in motion reveals mind in motion. *Frontiers in psychology* 2 (2011), 59.
[14] Tiziano Furlanetto, Cesare Bertone, and Cristina Becchio. 2013. The bilocated mind: new perspectives on self-localization and self-identification. *Frontiers in Human Neuroscience* 7 (2013). https://doi.org/10.3389/fnhum.2013.00071
[15] Shaun Gallagher. 2014. Phenomenology and embodied cognition. In *The Routledge handbook of embodied cognition*. Routledge, 9–18.
[16] Mar González-Franco, Daniel Pérez-Marcos, Bernhard Spanlang, and Mel Slater. 2010. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *2010 IEEE Virtual Reality Conference (VR)*. IEEE, Boston, MA, USA, 111–114. https://doi.org/10.1109/VR.2010.5444805
[17] Geoffrey Gorisse, Olivier Christmann, Etienne Armand Amato, and Simon Richir. 2017. First- and Third-Person Perspectives in Immersive Virtual Environments: Presence and Performance Analysis of Embodied Users. *Frontiers in Robotics and AI* 4 (2017). https://doi.org/10.3389/frobt.2017.00033
[18] Michael S. A. Graziano. 1999. Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position. *Proceedings of the National Academy of Sciences* 96, 18 (1999), 10418–10421. https://doi.org/10.1073/pnas.96.18.10418 arXiv:https://www.pnas.org/doi/pdf/10.1073/pnas.96.18.10418
[19] Yves Guiard. 1988. The Kinematic Chain as a Model for Human Asymmetrical Bimanual Cooperation. In *Cognition and Action in Skilled Behaviour*, Ann M. Colley and John R. Beech (Eds.). Advances in Psychology, Vol. 55. North-Holland, 205–228. https://doi.org/10.1016/S0166-4115(08)60623-8
[20] Lukas Heydrich, Trevor Dodds, Jane Aspell, Bruno Herbelin, Heinrich Buelthoff, Betty Mohler, and Olaf Blanke. 2013. Visual capture and the experience of having two bodies – Evidence from two different virtual reality techniques. *Frontiers in Psychology* 4 (2013). https://doi.org/10.3389/fpsyg.2013.00946

[21] Ken Hinckley, Randy Pausch, and Dennis Proffitt. 1997. Attention and Visual Feedback: The Bimanual Frame of Reference. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics* (Providence, Rhode Island, USA) *(I3D '97)*. Association for Computing Machinery, New York, NY, USA, 121–ff. https://doi.org/10.1145/253284.253318

[22] Ken Hinckley, Randy Pausch, Dennis Proffitt, and Neal F. Kassell. 1998. Two-Handed Virtual Manipulation. *ACM Trans. Comput.-Hum. Interact.* 5, 3 (sep 1998), 260–302. https://doi.org/10.1145/292834.292849

[23] Ken Hinckley, Randy Pausch, Dennis Proffitt, James Patten, and Neal Kassell. 1997. Cooperative Bimanual Action. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '97)*. Association for Computing Machinery, New York, NY, USA, 27–34. https://doi.org/10.1145/258549.258571

[24] J. A. Scott Kelso. 1995. *Dynamic patterns: The self-organization of brain and behavior.* The MIT Press, Cambridge, MA, US.

[25] J. A. Scott Kelso. 2021. The Haken–Kelso–Bunz (HKB) model: from matter to movement to mind. *Biological Cybernetics* 115, 4 (01 Aug 2021), 305–322. https://doi.org/10.1007/s00422-021-00890-w

[26] Scott R. Klemmer, Björn Hartmann, and Leila Takayama. 2006. How Bodies Matter: Five Themes for Interaction Design. In *Proceedings of the 6th Conference on Designing Interactive Systems* (University Park, PA, USA) *(DIS '06)*. Association for Computing Machinery, New York, NY, USA, 140–149. https://doi.org/10.1145/1142405.1142429

[27] Ryota Kondo and Maki Sugimoto. 2022. Effects of Body Duplication and Split on Body Schema. In *Augmented Humans 2022* (Kashiwa, Chiba, Japan) *(AHs 2022)*. Association for Computing Machinery, New York, NY, USA, 320–322. https://doi.org/10.1145/3519391.3524177

[28] Jaron Lanier. 2006. *Homuncular Flexibility.* Edge Foundation, Inc. https://www.edge.org/response-detail/11182

[29] Trevor Lee-Miller, Marco Santello, and Andrew M Gordon. 2019. Hand forces and placement are modulated and covary during anticipatory control of bimanual manipulation. *Journal of neurophysiology* 121, 6 (2019), 2276–2290.

[30] Bigna Lenggenhager, Michael Mouthon, and Olaf Blanke. 2009. Spatial aspects of bodily self-consciousness. *Consciousness and Cognition* 18, 1 (2009), 110–117. https://doi.org/10.1016/j.concog.2008.11.003

[31] Bigna Lenggenhager, Tej Tadi, Thomas Metzinger, and Olaf Blanke. 2007. Video Ergo Sum: Manipulating Bodily Self-Consciousness. *Science* 317, 5841 (2007), 1096–1099. https://doi.org/10.1126/science.1143439 arXiv:https://www.science.org/doi/pdf/10.1126/science.1143439

[32] Michal Luria, Samantha Reig, Xiang Zhi Tan, Aaron Steinfeld, Jodi Forlizzi, and John Zimmerman. 2019. Re-Embodiment and Co-Embodiment: Exploration of Social Presence for Robots and Conversational Agents. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19)*. Association for Computing Machinery, New York, NY, USA, 633–644. https://doi.org/10.1145/3322276.3322340

[33] Tamar R. Makin, Nicholas P. Holmes, and H. Henrik Ehrsson. 2008. On the other hand: Dummy hands and peripersonal space. *Behavioural Brain Research* 191, 1 (2008), 1–10. https://doi.org/10.1016/j.bbr.2008.02.041

[34] Tamar R. Makin, Nicholas P. Holmes, and Ehud Zohary. 2007. Is That Near My Hand? Multisensory Representation of Peripersonal Space in Human Intraparietal Sulcus. *Journal of Neuroscience* 27, 4 (2007), 731–740. https://doi.org/10.1523/JNEUROSCI.3653-06.2007

[35] Sebastian Marwecki, Andrew Wilson, Eyal Ofek, Mar Gonzalez Franco, and Christian Holz. 2019. Mise-Unseen: Using Eye Tracking to Hide Virtual Reality Scene Changes in Plain Sight. In *Proceedings of the 32nd Annual ACM Symposium on user interface software and technology (UIST '19)*. ACM, USA, 777–789. https://doi.org/10.1145/3332165.3347919

[36] Antonella Maselli and Mel Slater. 2014. Sliding perspectives: dissociating ownership from self-location during full body illusions in virtual reality. *Frontiers in Human Neuroscience* 8 (2014). https://doi.org/10.3389/fnhum.2014.00693

[37] Franz Mechsner. 2004. Perceptual-cognitive control of bimanual coordination. *International Journal of Sport and Exercise Psychology* 2, 3 (2004), 210–238. https://doi.org/10.1080/1612197X.2004.9671743 arXiv:https://doi.org/10.1080/1612197X.2004.9671743

[38] Mark R. Mine, Frederick P. Brooks, and Carlo H. Sequin. 1997. Moving Objects in Space: Exploiting Proprioception in Virtual-Environment Interaction. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '97)*. ACM Press/Addison-Wesley Publishing Co., USA, 19–26. https://doi.org/10.1145/258734.258747

[39] Reiji Miura, Shunichi Kasahara, Michiteru Kitazaki, Adrien Verhulst, Masahiko Inami, and Maki Sugimoto. 2021. MultiSoma: Distributed Embodiment with Synchronized Behavior and Perception. In *Augmented Humans Conference 2021* (Rovaniemi, Finland) *(AHs'21)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3458709.3458878

[40] Estelle Nakul, Nicolas Orlando-Dessaints, Bigna Lenggenhager, and Christophe Lopez. 2020. Measuring perceived self-location in virtual reality. *Scientific Reports* 10, 1 (22 Apr 2020), 6802. https://doi.org/10.1038/s41598-020-63643-y

[41] Albert Newen, Leon De Bruin, and Shaun Gallagher. 2018. *The Oxford handbook of 4E cognition.* Oxford University Press.

[42] 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

[43] Valeria I. Petkova and H. Henrik Ehrsson. 2008. If I Were You: Perceptual Illusion of Body Swapping. *PLOS ONE* 3, 12 (12 2008), 1–9. https://doi.org/10.1371/journal.pone.0003832

[44] Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa. 1996. The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology* (Seattle, Washington, USA) *(UIST '96)*. Association for Computing Machinery, New York, NY, USA, 79–80. https://doi.org/10.1145/237091.237102

[45] Jef Raskin. 2000. *The Humane Interface: New Directions for Designing Interactive Systems.* ACM Press/Addison-Wesley Publishing Co., USA.

[46] RootMotion. 2022. *RootMotion.* RootMotion. http://root-motion.com/

[47] Young June Sah, Minjin Rheu, and Rabindra Ratan. 2021. Avatar-User Bond as Meta-Cognitive Experience: Explicating Identification and Embodiment as Cognitive Fluency. *Frontiers in Psychology* 12 (2021), 695358.

[48] Jonas Schjerlund, Kasper Hornbæk, and Joanna Bergström. 2021. Ninja Hands: Using Many Hands to Improve Target Selection in VR. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 130, 14 pages. https://doi.org/10.1145/3411764.3445759

[49] Jonas Schjerlund, Kasper Hornbæk, and Joanna Bergström. 2022. OVRlap: Perceiving Multiple Locations Simultaneously to Improve Interaction in VR. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 355, 13 pages. https://doi.org/10.1145/3491102.3501873

[50] Dong-Hee Shin. 2017. The role of affordance in the experience of virtual reality learning: Technological and affective affordances in virtual reality. *Telematics and Informatics* 34, 8 (2017), 1826–1836.

[51] Daniel J Simons and Christopher F Chabris. 1999. Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events. *Perception* 28, 9 (1999), 1059–1074. https://doi.org/10.1068/p281059 arXiv:https://doi.org/10.1068/p281059 PMID: 10694957.

[52] Mel Slater, Daniel Pérez Marcos, Henrik Ehrsson, and Maria Sanchez-Vives. 2009. Inducing illusory ownership of a virtual body. *Frontiers in Neuroscience* 3 (2009). https://doi.org/10.3389/neuro.01.029.2009

[53] Mel Slater, Bernhard Spanlang, Maria V. Sanchez-Vives, and Olaf Blanke. 2010. First Person Experience of Body Transfer in Virtual Reality. *PLOS ONE* 5, 5 (05 2010), 1–9. https://doi.org/10.1371/journal.pone.0010564

[54] Nigel Stepp and Michael T Turvey. 2010. On strong anticipation. *Cognitive systems research* 11, 2 (2010), 148–164.

[55] Kazuma Takada, Midori Kawaguchi, Akira Uehara, Yukiya Nakanishi, Mark Armstrong, Adrien Verhulst, Kouta Minamizawa, and Shunichi Kasahara. 2022. Parallel Ping-Pong: Exploring Parallel Embodiment through Multiple Bodies by a Single User. In *Augmented Humans 2022* (Kashiwa, Chiba, Japan) *(AHs 2022)*. Association for Computing Machinery, New York, NY, USA, 121–130. https://doi.org/10.1145/3519391.3519408

[56] Adrien Verhulst, Yasuko Namikawa, and Shunichi Kasahara. 2022. Demonstrating Parallel Adaptation: How Switching between Two Virtual Bodies with Different Perspectives Enables Dual Motor Adaptation. In *SIGGRAPH Asia 2022 XR* (Daegu, Republic of Korea) *(SA '22)*. Association for Computing Machinery, New York, NY, USA, Article 4, 2 pages. https://doi.org/10.1145/3550472.3558416

[57] Andrea Stevenson Won, Jeremy Bailenson, Jimmy Lee, and Jaron Lanier. 2015. Homuncular Flexibility in Virtual Reality. *Journal of Computer-Mediated Communication* 20, 3 (2015), 241–259. https://doi.org/10.1111/jcc4.12107

[58] Alfred L. Yarbus. 1967. *Eye Movements and Vision.*