Master's Thesis

Embodiment and the Proteus Effect in AR: Perspective, Avatar Appearance, and Transition

Riku Otono

March 15, 2023

Graduate School of Information Science
Nara Institute of Science and Technology

A Master's Thesis
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
MASTER of SCIENCE

Riku Otono

Thesis Committee:

Professor Kiyoshi Kiyokawa	(Supervisor)
Professor Hirokazu Kato	(Co-supervisor)
Associate Professor Hideaki Uchiyama	(Co-supervisor)
Assistant Professor Naoya Isoyama	(Co-supervisor)
Assistant Professor Monica Perusquia-Hernandez	(Co-supervisor)

Embodiment and the Proteus Effect in AR: Perspective, Avatar Appearance, and Transition*

Riku Otono

Abstract

In this thesis, I explore how avatar embodiment in Augmented Reality (AR) impacts the Proteus effect. This is done through three studies with different experiment conditions. In the first study, I compare different perspectives (first person perspective (1PP) and third person perspectives (3PP)), different avatar age (young and elderly) and the impact on the Proteus effect in AR. In the second study, I investigate how two virtual avatars with different body shapes (non-muscular and muscular) affect users' physical performance. In the third study, I examine the impact of visual transitions from a regular avatar to a muscular avatar.

AR embodiment has little been explored, compared to Virtual Reality (VR), due to some technical challenges. Unlike VR environments, a user own body is visible in the real environment. In addition, tracking error and latency might deteriorate embodiment experience especially when the avatar body has different shape from the user real body. This is especially the case with 1PP AR embodiment, where the user can see their own real body with an AR avatar overlay from egocentric view. To deal with these technical challenges, I implemented a 3PP AR embodiment system which allow a user to observe the avatar from behind. In the presented system, a user real body was blurred (study #1) or completely erased (study #2 and #3).

*Master's Thesis, Graduate School of Information Science,
Nara Institute of Science and Technology, March 15, 2023.

Furthermore, I employed several methods to examine the Proteus effect. In study #1, I measured the smoothness of walking and the time to walk in a certain distance to confirm the Proteus effect of an elder avatar. In study #2, I counted the number of lifting a dumbbell and hand-grip strength to measure the user's physical performance. In study #3, to evaluate the impact of the Proteus effect of a muscular-shaped avatar, I assessed the surface electromyography (sEMG) magnitude.

As a result, I confirmed that the 3PP decreases the smoothness of walking during embodiment in study #1. In study #2, I discovered the muscular-shaped avatar appearance supports the immediate and prolonged Proteus effect. Finally, through study #3, I found that changing an avatar's appearance with visual feedback has a positive impact on retaining the sense of agency. In addition, the results showed that the impact of the immediate Proteus effect was boosted by active transitions with visual feedback.

To conclude, the presented system is a promising tool to enable users to feel embodiment in the real environment and receive the strong Proteus effect, especially using a muscular avatar. The demonstrated phenomena might be used in AR training for sports employing a muscular-shaped avatar. Further, the impact of the Proteus effect for a muscular-shaped avatar could be applied to other areas which need physical tasks such as construction and logistics.

In the future, I'll study an avatar's appearance change using an avatar with a different appearance such as a taller avatar. Furthermore, it will be important that future research investigate the impact of changing an avatar's appearance using other types of physiological sensors than sEMG sensors, such as electroencephalogram (EEG) and electrodermal activity (EDA).

Keywords:

Avatar, Proteus effect, Perspective, Transition, Augmented reality

Contents

List of Figures

1 Introduction

When people experience Virtual Reality (VR) or Augmented Reality (AR) they are often represented as three-dimensional virtual avatars. Virtual avatars are especially featured in social VR and AR applications, such as VRChat*, Meta's Horizon Worlds†, Spatial‡, and others. For example, VRChat has 7.3 million total users according to MMO Stats§. These large number of users increase the importance of research into how to make users feel that their avatars are their own body and how this sense of embodiment can benefit the users.

A previous VR study has identified a large number of factors that may strengthen the sense of embodiment such as animation techniques, point of view, or avatar appearance [14]. Strong embodiment experiences can lead the so-called Proteus effect. This was defined as the effect of avatars' appearance on the user behavior [91]. This effect causes changes in user self-perception and behavior, sometimes improving the user's performance [91]. For instance, embodying a taller avatar may lead to the user behaving more confidently in a negotiation task [91]. The study by Yee and Bailenson was followed by many VR studies that I elaborate in the chapter 2.

Avatar embodiment and the Proteus effect have been investigated using AR environment [70, 88, 89]. The Proteus effect in AR and the physical world could ameliorate people's lives in a variety of domains such as fitness, education, and the clinical fields. However, some technical challenges around AR embodiment still remain [16]. First unlike VR, the user's real body is visible in AR, especially when the avatar body structure is significantly different from the user's body shape [16]. Registering the virtual avatar body correctly on the user's cor-

*https://hello.vrchat.com/
†https://www.oculus.com/horizon-worlds/
‡https://www.spatial.io/
§https://mmostats.com/game/vrchat

responding real body is also challenging in a first-person perspective (1PP) due to tracking errors [16].

However, these challenges are expected to be overcome by future advanced technologies. For instance, video inpainting methods have been drastically improved thanks to better deep learning techniques [30,31,35,38,46]. These inpainting techniques could be applied to completely erase a visible user body in real-time. Furthermore, motion capture from an egocentric view has been addressed [6, 65, 90]. These novel techniques would allow stable AR avatar embodiment with flawless tracking.

In this thesis, I explore how AR avatar embodiment impacts the Proteus effect. This is done through three studies with different experiment conditions. In the first study, I compared different perspectives (first person perspection (1PP) and third person perspectives (3PP)) and different avatar age (young and elderly) and the impact on the Proteus effect in AR. In the experiment, I used the system which blurs a user's real body to let them focus on the avatar's appearance. As a result, I found that a 3PP induce as much embodiment as a 1PP. However, the results did not support the Proteus effect. Then, I decided to completely erase the real body to ease avatar embodiment in the following study since I found negative feedback about the blurred body image from some participants. Therefore, in the second study I investigated how two virtual avatars having different body shapes (normal and muscular) affect user physical performance. In the user study in the second study, I implemented a 3PP AR embodiment system which completely erases a user's real body. The results showed the Proteus effect in the 3PP AR. In the third study, I devised the visual transition to let a user smoothly embody in the avatar and enhance the Proteus effect in AR. In the user study, I examined the impact of visual feedback and physical action during transition from a regular-shaped avatar to a muscular-shaped avatar. the results demonstrated that active transition with visual feedback help maintain the component of embodiment and partially enhance the Proteus effect while a user embody in the avatar.

In the rest of the thesis I first review related work and highlight the novelty of our research (Chapter 2). In Chapter 3 I evaluated the Proteus effect in a 1PP and a 3PP AR with blurred user body. In Chapter 4 I employed a fully erased user body and examined the Proteus effect in AR. In Chapter 5 I explored avatar transitions in AR. I discussed our findings in Chapter 6 followed by a Conclusion section with directions for future work (Chapter 7).

2 Related work

This chapter details the previous work on embodiment, the Proteus effect, and visual transitions that inspired my work and describes how AR and VR has been employed to improve user physical performance through avatars.

2.1 The Sense of Embodiment

Extensive previous work has shown that it is possible to experience an SoE toward another body than our own, referred to as a body ownership illusion. The most famous example illustrating this is perhaps the rubber hand illusion, during which a rubber hand is perceived as one's real hand thanks to synchronous visuotactile stimulation [3, 40, 84, 85].This illusion was shown to work in VR environments using VR hands [61, 68, 71, 81–83, 92], and eventually with full-bodied avatars [5, 29, 42, 58].

Many studies have attempted to improve the understanding of how these illusions work through the assessment of the SoE [14, 17, 18, 28, 50]. They generally focused on three sub-components of this sense: (1) the sense of agency, (2) the sense of ownership, and (3) the sense of self-location [28]. These studies revealed that multiple inter-dependent factors could impact these individual senses including the avatar's appearance, degree of control, and embodiment point of view [14]. In particular, some studies on the avatar's appearance have found that it was preferable to use cartoonish or non-human avatars compared to realistic human ones due to the Uncanny Valley effects that they cause [41]. However, several recent papers contrast with these results and suggest a more intricate relationship with the SoE: in these studies, avatars that differ a lot from the appearance of their users tended to result in weaker body ownership illusions [11, 24, 51, 86].

Different aspects of these three sub-components have been explored in previious studies, including the visuo-motor stimulations [87], the visuo-tactile stimulations [50], and the avatar's appearance [18, 41, 55, 87]. The inter-relationships between avatar appearance, avatar control and user point of view have also been investigated [14]. They found that the point of view and avatar control levels were consistently prioritized by users compared to avatar appearance when it comes to enhancing the SoE [14].

Regarding the impact of the user's point of view, many VR researchers feel that the 1PP is superior to the 3PP in inducing the SoE [19, 22, 42–44, 57, 69, 72]. However, multiple VR studies showed that using a 3PP can also provide strong body ownership illusions [8, 15, 17, 18, 36, 37, 60]. Gorisse et al. (2017) notably found that the 3PP can provide the same amount of agency as the 1PP [17]. Other studies also found that the senses of ownership and self-location could be improved by the 3PP under multi-sensory congruence conditions where the user feels haptic feedback synchronized with a VR scene [36, 37, 60]. However, little research exists to confirm these results in AR.

A 3PP system was also used to investigate the SoE and its sub-components in an AR environment where the user observes their avatar in front of them [66] 2.1. This shows that consistency between avatar and user movements improves the sense of ownership [66].

The recent review of Genay et al. (2021) highlighted that using AR for body ownership illusions presents specific challenges that make the induction of the SoE within it particularly difficult [16]. These challenges include having to deal with the user's real body visibility, the contrasting rendering of real and virtual things, and the tracking errors and latency that are easier to notice than in VR, especially in the 1PP. As integrating the avatar into the real world cannot be done perfectly with current technologies yet, I had the feeling that AR setups might have even further interest in having visual transitions to mitigate the potential effects of discrepancies resulting from current system limits. This participated to my decision of focusing on AR.

finding that AR embodiment has some specific challenges, including real body visibility, tracking errors, and latency. This is especially the case with 1PP AR embodiment, where the user can see their own real body with an AR avatar overlay. Even with current technology, it is difficult to make the user's own body

Figure 2.1: The experimenter's view of a participant standing two meters behind the AR marker (left), the user's view of an anthropomorphic avatar (center), and a non-anthropomorphic avatar on the AR marker (right) [66]

invisible and precisely overlay the avatar body [13, 26]. I address this issue by blurring or erasing a user's own body.

2.2 The Proteus effect

The Proteus effect describes a phenomenon where users "conform to the behaviour that they believe others would expect them to have" while embodying an avatar [91]. This effect has been intensively studied in VR since Yee and Bailenson coined the term in 2007. A meta-analysis of the Proteus effect found that it occurs at different levels [63] and that self-similarity increases the effect [53]. On the other hand, it was rarely studied in AR due to the technical challenges mentioned previously.

Several VR papers have investigated using the Proteus effect to increase the user's physical strength [32]. For example, Banakou et al. (2018) found that embodying Albert Einstein in 1PP VR could momentarily improve the user's cognitive abilities [1]. Kocur et al. (2020) found that participants embodying muscular-shaped avatars had a lower perceived exertion during an isometric force task [33]. They also found that such avatars could influence the user's heart

Figure 2.2: User's view from a third-person perspective in Phantact [70].

rate [32]. Osimo et al. (2015) showed that the Proteus effect led to cognitive changes by using an avatar that looks like Sigmund Freud [52]. Lin et al. (2021) further found that such effects can be obtained during actual physical exercises even after the embodiment of the muscular-shaped avatar ended [39]. The effect of embodiment time on the impact of the Proteus effect was also proposed by hypothesising that the longer users embody an elderly avatar, the slower they will walk in VR [34].

The Proteus effect can be applied in several domains. For instance, it was used to address the fear of heights using a bird avatar [53], and led to cognitive changes by using an avatar looks Dr Sigmund Freud [52]. One aspect of these effects is that they can be prolonged even after the virtual embodiment has ended. In psychological studies, the Proteus effect affected the participants' behaviors after they finished their assigned tasks [64,67]. Rosenberg et al. (2013) investigated whether participants helped other people or not after embodying superheroes. They did this by counting how many spilled pens they picked up for the experimenter, and participants embodied as superhero avatars picked up

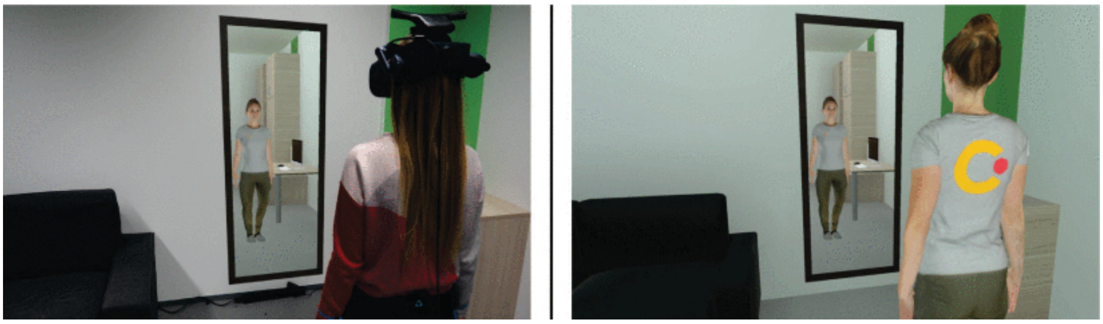

Figure 2.3: A user viewing their avatar reflected in the mirror in AR (left) and in VR (right) [88]

significantly more pens than the control group [67]. Reinhard et al. (2020) also showed that participants previously embodied in an older avatar required more time to walk the same distance after the embodiment than those previously embodied in younger avatars [64]. However, the Proteus effect occurred only in the first session after the stimulus and not in the second session [64]. The effect of embodiment time on the impact of the Proteus effect was also proposed by hypothesising that the longer users embody an elderly avatar, the slower they will walk in VR [34].

The Proteus effect in AR has also been explored, but less extensively. For example, Wolf et al. (2020) used a video see-through HMD to investigate how users perceive their avatars observed in a virtual mirror in AR compared to in VR settings [88]. They found that using an AR see-through HMD can provide a similar level of SoE as a fully immersive VR HMD, although presence partially decreased [88]. A current study compared optical and video see-through displays and found that both systems supported similar effects of avatars [89].

In this work, I examine the Proteus effect by evaluating the user's physical performance. In the study #1 I used a elder avatar to check how an avatar's age affects the user's physical performance. In the study #2 and the study #3 I replicate previous work demonstrating that a muscular avatar can improve physical performance to test if this performance varies when applying different types of avatar transitions.

2.3 Visual transitions

Several types of transitions for avatars have been investigated in the past. In VR research, transitions were previously designed to provide a smooth experience when entering an immersive virtual environment [25, 75]. For example, Sproll et al. (2013) created a transition method staged in five steps for the user to switch from their own body in the real environment to being inside an avatar's body within a virtual environment [75]. Previous studies that tested transitions between a virtual replica of the real environment and a virtual world showed improved subjective user experience and sense of presence [73, 76–78]. Smolentsev et al. (2017) also investigated similar transitions using a 2D monitor and demonstrated that transitions from a virtual replica increased spatial presence [74]. Jeffrey et al. (1999) showed a transition effect from the real to a virtual environment using a stereo camera [59].

Transitions were also previously implemented to switch from one virtual body to another [23, 79]. Dynamic morphing can be used to apply changes to virtual content over time in a smooth way. When slow enough, the changes can even go unnoticed by the user. This is of particular interest when trying to maintain a continuous, coherent experience of avatar embodiment, as shown by Jung et al. (2020) [23].

Lastly, previous research explored the use of transitions to change the user's visualization perspective from the first-person perspective (1PP) to the third-person perspective (3PP). Such kind of change is valued in applications like video games to extend the user's spatial awareness of their surroundings. Indeed, flexibility in shifting perspectives was shown to increase the user's relationship with the character [9]. In VR games, the 1PP helps players to focus on in-game action and enhances involvement, whereas the 3PP shifts their focus to the player character and game interaction [12]. Transitioning between these two perspectives allows users to benefit from both points of view without being disoriented [7, 15, 22]. Debarba et al. (2017) tested three approaches using different types of camera movements and found that very fast linear translation of the camera with a slightly blurred vision was optimal in terms of user experience and cybersickness [15]. In another study, users could change their perspective in VR and it was found that the ratings of presence, agency, and embodiment were lower in the 3PP although benefits of the 3PP were also found [22]. However, none of the transitions mentioned above was evaluated within an AR context and little is known about their effects on the SoE.

In the study #3, I combine several types of transitions and evaluate the effects of using them to switch of avatar on the SoE in AR for the first time.

2.4 Summary

As can be seen from the related work, there has been significant research conducted in the embodiment in VR. This includes research in avatar appearance, perspectives, and avatar body morphing. However there has been little research to explore these same effects in AR. My thesis addresses this research gaps, and specifically makes the following significant contributions:

1. It investigates the effect of perspective and avatar appearance in AR on user SoE and physical performance.
2. It examines the Proteus effect of a muscular avatar under fully erased user body conditions.
3. It illustrates how the avatar transitions influence the Proteus effect in AR.

3 Study #1: the impact of perspective and avatar age under blurred user body

3.1 Overview

This chapter reports on a user study that investigated the Proteus effect of avatar age for different perspective conditions in Augmented Reality (AR). In the experiment, I blurred a user's real body to let them focus on the avatar's appearance in the 3PP. As I mentioned in section 2, previous study showed the Proteus effect of an elder looking avatar on users walking speed [64]. Our goal is to confirm that such Proteus effect can be induced in the 3PP AR with a blurred real body.

I employ an AR system, which allows users to embody toward their avatars either from the 1PP or the 3PP, called Phantact [70]. Using the system, I compare performance with the 1PP and the 3PP in AR on a walking task. In the 1PP, the user can see a pre-captured 360 image representing the real environment and look down the avatar body through the avatar's eyes (see Fig. 3.1). In the 3PP, the user's real body is blurred and the avatar body is overlaid onto it (see Fig. 3.2). In both perspectives, the virtual avatar body moves in sync with the user's real body movement.

It is not clear how different body perspectives affect the SoE and the Proteus effect in AR. In the real world, people always see the world from the 1PP. When people apply the 1PP to embody in avatars, generally they have to use a virtual mirror to check their avatars' whole appearance. Therefore, it might be difficult for users to keep feeling embodiment toward their avatars when a virtual mirror is removed.

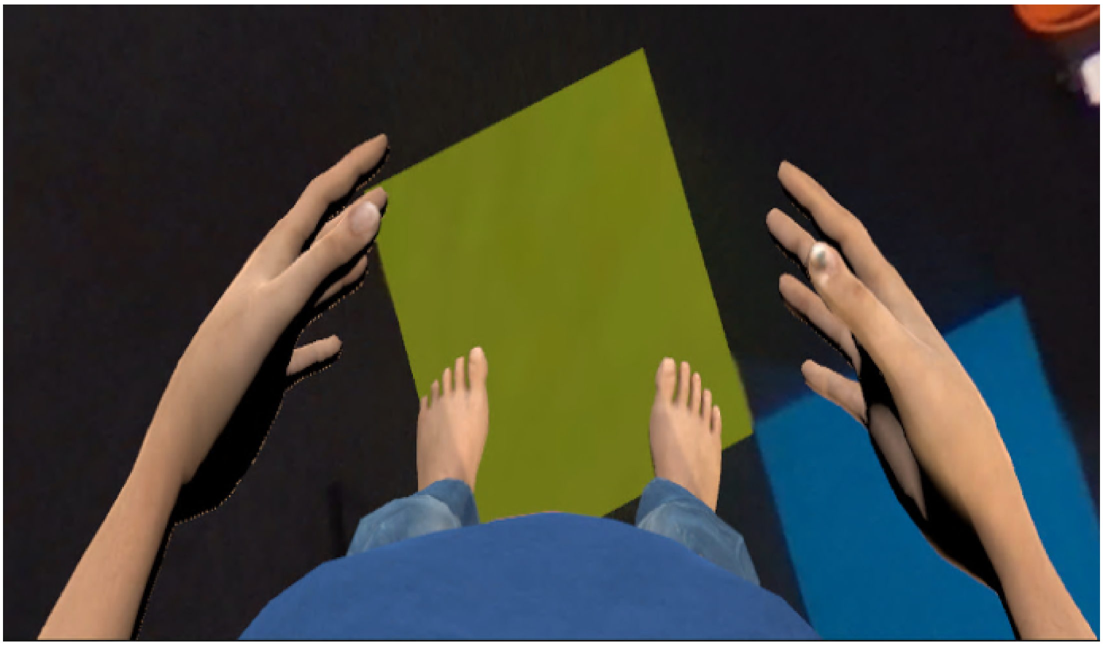

Figure 3.1: User's view from the 1PP in the Phantact [70].

On the other hand, the 3PP allows users to continuously see their avatars' appearance without a virtual mirror. Therefore, I hypothesised that the 3PP can induce as strong SoE and the Proteus effect during embodiment as the 1PP in AR environments without a virtual mirror. Furthermore, I expected that the Proteus effect would be prolonged longer after the 3PP embodiment than the 1PP because the 3PP offers the user more opportunities to remind that they are embodied in a virtual avatar.

My hypotheses are summarized as below:

H1 The intensity of the sense of embodiment from the 3PP is similar to that from the 1PP.

H2 The Proteus effect is induced in both the 1PP and the 3PP.

H3 The Proteus effect is induced longer in the 3PP than in the 1PP.

To explore these hypotheses, a user study was conducted comparing the sense of embodiment **H1** measured by self-report after the task. Plus, to examine the Proteus effect of an elder avatar, I focused on the walking performance during and

Figure 3.2: User's view from the 3PP in Phantact [70].

after embodiment. Generally, people have a stereotype that an elder people walk roughly and slowly compared to young people. Therefore, I measured smoothness of walk-in-place performance **H2**, and post-hoc walking time **H3**. Through the experiment, it was confirmed that there was no significant difference in the SoE between the 1PP and the 3PP (confirming **H1**), the immediate and the prolonged Proteus effect was not evident (**H2** and **H3** were not supported). In addition, I found that the 3PP decreases the smoothness of walk-in-place performance compared to the 1PP.

3.2 Methodology

In this study, I aimed to figure out the effects of perspectives and avatar appearance in AR. To do that, I implemented an AR embodiment system based on the

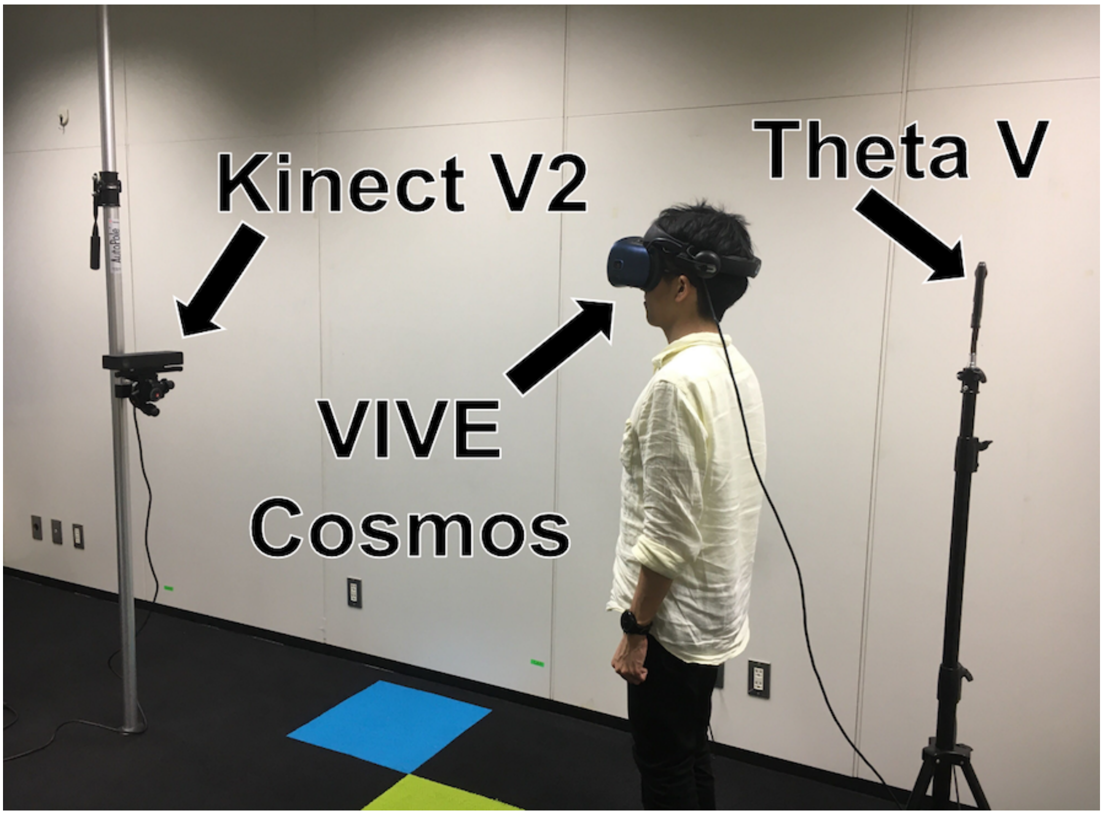

Figure 3.3: Hardware configurations in the experiment.

3PP AR embodiment system (Phantact [70]). In the implemented system, users can embody in their avatars either in the 1PP and the 3PP.

3.2.1 First-person perspective

The view from the 1PP was made using a 360 degree panorama still image captured by the Ricoh Theta V. The panorama image was taken before the experiment from 1.65 meters above the ground at the specified standing position to simulate the participants' view. I chose a static image instead of real-time video streaming because it was difficult to precisely overlay an avatar on the user's body. This enabled a virtual avatar to be placed over the view of the real world without showing the user's real body (see Fig. 3.1).

3.2.2 Third-person perspective

In the 3PP, I used real-time 360 degree video streaming captured by the Ricoh Theta V. In our system, a camera stand was used to fix the camera, while in the original Phantact a backpack was used to mount the camera on the user [70]. Figure 4.4 illustrates the hardware configurations.

The synchronized avatar was overlaid on the users' body to induce the SoE in AR from the 3PP. To achieve this, the video was transferred through WebSockets to a second PC to extract the users' body region using the BodyPix* body segmentation library and then blur the body region. Then the processed video was transferred back to the PC and the synchronized avatar was overlaid on the video background. Figure 3.2 shows the result from the body segmentation and avatar overlay. The blurred body movement delayed around 1.0 second from the avatar movement due to processing time to segment and blur a human body region.

3.2.3 Implementation details

A prototype system was developed using the Unity game engine (2019.3.11) running on a PC with Windows 10 OS. In both the 1PP and 3PP, a Ricoh Theta V (1080p) 360 degree video camera was used to capture a 360 degree panorama view of the users surroundings.

In the prototype, the user wore the VIVE Cosmos head mounted display (HMD), which has a viewing angle of 110 degrees, 90 Hz refresh rate, and a resolution of 2880 × 1700 pixels (1440 × 1700 per eye). A Kinect motion sensor for Windows V2 was used to capture the user's body motion. The Kinect has a RGB camera with a resolution of 1920 × 1080 pixels, and a depth sensing infrared camera with a resolution of 512 × 424 pixels. The user's motion was reflected in the avatar using joint angles obtained by the Kinect.

*https://github.com/tensorflow/tfjs-models/tree/master/body-pix

3.3 User study

3.3.1 Experimental design

The experiment followed a 2 × 2 factorial design. The independent variables were the avatar representation (young/elderly) and the perspectives (1PP/3PP). Therefore, I had the following four conditions:

- **Condition Y_1PP**: a young avatar with the 1PP
- **Condition Y_3PP**: a young avatar with the 3PP
- **Condition E_1PP**: an elder avatar with the 1PP
- **Condition E_3PP**: an elder avatar with the 3PP

In each of these conditions the participants performed the following task. While embodying an avatar, participants were asked to walk for one minute. Due to the scale limitation of the laboratory, walking task was performed in a place. After embodiment, participants were instructed to walk to a table which was five meters away from the task place. I measured the time participants walked this distance without they realizing.

I used a within-subjects experiment design. So, the participants performed the task in all the conditions in a counterbalanced order based on a Latin-square design. Each participant selected one of the two avatars that were created by using MakeHuman[†] according to their claimed gender (see Fig.3.4).

3.3.2 Measures

I measured the ense of embodiment using the sense of embodiment questionnaire. I also evaluated the strength of the Proteus effect produced by an elder avatar by measuring the smoothness of walking, and post-hoc walking time. More information about both of these are described below.

[†]http://www.makehumancommunity.org/

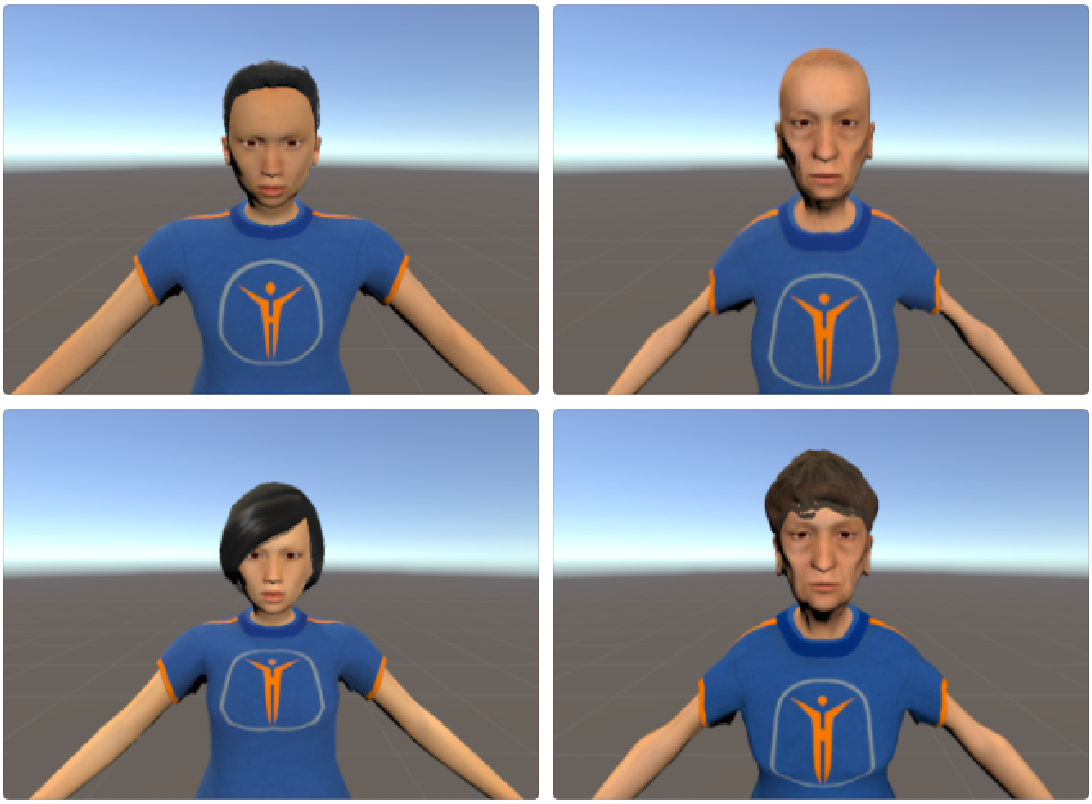

Figure 3.4: Appearance of the young male avatar (upper left), the elder male avatar (upper right), the young female avatar (lower left) and the elder female avatar (lower right).

SoE self-report

To explore hypothesis **H1**, I measured the SoE after each condition using the questionnaire proposed by [54]. I expected that the 3PP would offer similar SoE as the 1PP because the 3PP allows users to confirm the avatar body. This embodiment questionnaire consists of 16 questions with four interrelated subscales (Appearance, Response, Ownership, and Multi-Sensory). A 7-point Likert-scale was used in each question (where 1 = Strongly disagree and 7 = Strongly agree) and the embodiment score was calculated based on each score from the questions. The items of the questionnaire is listed in 9. The order of the questions was randomized to reduce the order effect.

Since we are using this questionnaire we could modify the first hypothesis to be:

H1 There will be no significant difference between the 1PP and the 3PP conditions in the sense of embodiment as measured by the Peck SoE questionnaire [54].

Smoothness of walking

To explore hypothesis **H2** and measure the Proteus effect, I measured the smoothness of walking of the experiment participants. This is because generally elderly people move less smoothly than young people, so if there is a strong Proteus effect with an elderly avatar then the user's movement should be less smooth.

To measure the smoothness of walking I used the jerk cost. The jerk cost refers to the mean square of differentiated values of acceleration, and is used to measure the smoothness of movements (higher jerk cost indicates less smooth movement) [21]. So with a strong Proteus effect, the jerk cost should be expected to be higher with the elder avatar than that with the young avatar. I was able to calculate jerk cost, using the following equation:

$$J = \frac{1}{n}\sum(f'''\frac{Left_k+Right_k}{2})^2 \tag{3.1}$$

where J is jerk cost, n is the number of samples,$Left_k$ is root mean square of x, y, and z coordinates of the left knee, and $right_k$ is root mean square of x, y, and z coordinates of the right knee. Joint coordinates were measured by the Kinect V2 at 30 fps.

I expected that if there was a strong Proteus effect, while embodied in an elder avatar, the smoothness of walking would decrease. Hence, the second hypothesis was modified to be:

H2 In terms of smoothness of walking, As measured by smoothness of walking, the Proteus effect of an elder avatar would be induced while embodiment both in the 1PP and the 3PP.

Post-hoc walking time

To explore hypothesis **H3**, the post-hoc walking time was also measured. Participants could observe their avatar appearance continuously even after removing the virtual mirror in the 3PP. On the other hand, participants could not observe the whole avatar body after removing the mirror. Therefore, the 3PP provides the user more chances to remind that they are embodied in a virtual avatar. Thus, in comparison to the 1PP, the Proteus effect would be prolonged. As a result, the post-hoc walking time in the 3PP was expected to be longer than that in the 1PP.

Unlike measurement of walking smoothness during embodiment, participants actually walked in post-hoc walking time measurement. To measure this, the method from the experiment in [64] was used. The walking distance was set from the task space to the table (five meters) and the walking time was measured using a web camera (Logitech C922 Pro HD Stream Webcam). Before the experiment, I decided pixel positions that correspond the start and end line of the walking distance. Based on the pixel values, I detected when participants cross the start and end line without participants realizing the measurement. Participants were asked to walk from the task space to the table, and to avoid any unwanted bias they were not informed that the walking time would be measured.

Considering the use of walking time, the third hypothesis was modified to be:

H3 In terms of the post-hoc walking time, the Proteus effect of an elder avatar would be induced longer in the 3PP than in the 1PP.

3.3.3 Procedure

When a person came to participate in the study, I first briefly explained our experiment. Next, I guided them to the table where they signed a consent form and filled in a pre-questionnaire about their background (e.g. experience with HMDs). Then, I led them to the task space that was approximately five meters away from the table. The participants put on the HMD and started one of the four conditions in the task space. In each condition, they first watched their self-avatar for one minute reflected in the virtual mirror. Figure 3.5 illustrates the male version of the participant's views in each condition while seeing the virtual mirror.

Participants could see the moving avatar body synchronized with their own body in either perspective, 1PP or 3PP. In the second part of the experiment, I

Figure 3.5: Participant's views with the young male avatar in the 1PP (upper left) and the 3PP (upper right), and those with the elder male avatar in the 1PP (lower left) and the 3PP (lower right).

asked them to walk-in-place for one minute. They took off the HMD after the walking task and went back to the table to fill in the questionnaire.

3.3.4 Participants

A total of eight participants were recruited from the university (four females, four males) with their ages ranging from 22 to 37 years old ($M = 25.75$, $SD = 4.74$). All participants had experienced AR/VR at least once.

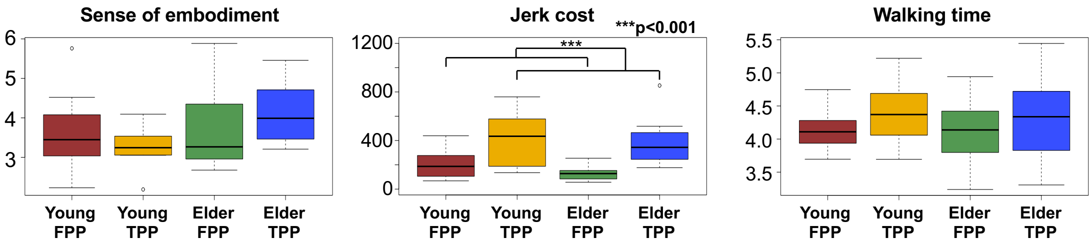

Figure 3.6: Results of the sense of embodiment (left), jerk cost (middle), and walking time (right).

3.4 Results

3.4.1 SoE self-report results

To analyze the SoE questionnaire results, I treated the resulting scores as non-parametric data and applied the aligned rank transform (ART). I conducted an ANOVA for the transformed data to verify **H1**. The result (see Fig.3.6(left)) shows, as expected, that there was no significant difference between 1PP and 3PP ($F(1,28)=0.001$, $p=0.981$), so **H1** is supported.

3.4.2 Smoothness of walking

The jerk cost was calculated to test **H2**, and the average jerk cost results for elderly and young avatars are shown in Figure 3.6. The result of a two-way ANOVA (see Fig.3.6(middle)) shows no significant difference in the jerk cost between young avatars and elder avatars ($F(1,28)=0.606$, $p=0.443$), meaning that the Proteus effect was not evident from this measure. On the other hand, I could confirm a significant difference in the jerk cost between the different perspectives ($F(1,28)=14.606$, $p=0.001$). This result indicates that the 3PP decreases smoothness of walking compared to walking in the 1PP. I speculate that the participants were not used to using the 3PP, resulting in lower smoothness of walking. Thus, **H2** is rejected.

3.4.3 Post-hoc walking time

Figure 3.6 shows the average walking time for participants in each condition. I used a two-way ANOVA to test **H3**. Unfortunately, the result shows that the Proteus effect was not confirmed in post-embodiment performance from these measures between avatar age ($F(1,28)=0.080$, $p=0.780$) and perspectives ($F(1,28)=1.569$, $p=0.221$). Thus, **H3** is rejected.

3.5 Discussion

In the following, I discuss the reasons why the results support **H1** and why **H2** and **H3** were rejected.

Both in the 1PP and 3PP conditions, participants could see their avatar bodies reflected by the virtual mirror. In addition, they could confirm that the avatar bodies moved in sync with their real-time body movement. Therefore, the SoE was induced, as a previous AR study showed a visuo-motor synchronization between the user and the avatar induces a SoE [66].

However, the tracking of some parts of body was not precise, because I only used the Kinect for body tracking. For instance, I got some negative comments from participants related to hand tracking. For example participant P3 said "the finger movement was not precise", and participants P7 said "the avatar sometimes looked jittering". Consequently, participants might not have been able to feel enough embodiment to induce the Proteus effect. This may be one explanation for why **H2** was rejected.

Participants could see their own blurred body in the 3PP. However, there was a communication delay between the two PCs and image processing time was needed to perform body detection and blurring, caused a delay. This meant that the the motion of the blurred body appeared slower than that of the user's avatar. As a result, the participant might not have been able to concentrate on their avatar, and were distracted by the blurred, delayed image of their own body. For example participant P3 said "a real body (blurred) was noticeable and felt like watching two bodies (a real body and an avatar body) at the same time". This may be also have affected the result for **H2**.

The distance of the post-hoc walking from the task space to the table may have been too short. Due to the limitation of space, I only could set the distance to 5 meters, which is less than the length used by Reinhard et al. [64]. This may be one explanation for why **H3** was rejected.

3.5.1 Experiment limitations

Although the results of the user study were interesting, there are some limitations. First, the user view in the 1PP was not real-time because a static 360-image was used in the 1PP while a real-time 360 video was displayed in the 3PP. In a real-time scene, the user's own body is visible and should be erased, and the avatar body should be overlaid on the user's body region. However, segmenting body region from the 1PP in real-time is challenging even with state of the art methods, and it was also difficult to overlay the avatar body on the user body region. As a result, I decided to use a static image in the 1PP. In the future this limitation could be overcome by using more advanced image segmentation and in-filling techniques.

Secondly, the post-hoc walking time was measured just after the task, thus it is not clear how long the Proteus effect was prolonged for. Previous study measured twice the walking speed after embodiment, and they found that the Proteus effect diminished quickly as the effect occurred in the first session and not in the second session [64]. By measuring the performance several times after embodiment, I would figure out how the Proteus effect attenuates after embodying in an avatar.

Furthermore, I only had 8 participants, which might have been too few to get significant difference. It would be good to have a larger number in the future.

3.6 Conclusion

In this paper, I investigated the Proteus effect of avatar age for 1PP and 3PP conditions in AR. I developed a prototype AR system that allowed people to see themselves as elderly or young avatars, and then conducted an experiment to verify the following three hypotheses.

H1 There will be no significant difference between 1PP and 3PP conditions in the sense of embodiment as measured by the Peck SoE questionnaire.

H2 In terms of smoothness of walking, the Proteus effect of an elder avatar would be induced while embodiment both in the 1PP and the 3PP.

H3 In terms of the post-hoc walking time, the Proteus effect of an elder avatar would be induced longer in the 3PP than in the 1PP.

Using Peck's SoE questionnaire [54] it was confirmed that the 3PP can produce as much sense of embodiment as the 1PP. However, the Proteus effect during-embodiment performance and post-embodiment performance was not confirmed from the jerk cost and walking time although. So hypothesis H1 was confirmed, while H2 and H3 was rejected. It was also confirmed that 3PP, compared to 1PP, decreases the smoothness of walking. The reason for this is not clear, and should be studied in future work.

There were also some limitations with the work. For example, a static image was used in the 1PP while a real-time video streaming was displayed in the 3PP. I should overcome some technical challenges such as body tracking and image processing to induce the Proteus effect more clearly in AR.

In the future, I would like to conduct follow-up experiments based on the lessons learned from this study, and compare 1PP and 3PP about the persistence of the Proteus effect in more detail. To address that, there is a need to overcome the technical challenges of the 1PP, erasing the user's own body and overlaying the avatar body. Furthermore, the post-embodiment Proteus effect has to be explored by measuring the effect for a longer time after the task. For example in [39] the effect was measured the day after the task.

4 Study #2: the impact of avatar musculature under fully erased user body

4.1 Overview

In study #1, I investigated the impact of perspective and avatar age in 3PP and 1PP AR environments. In AR, unlike VR, the user's real body is visible in the real environment which could distract users from focusing on their virtual avatar. To reduce the effect of this, in study #1 I overlaid the user's virtual avatar body onto a blurred image of their real body. However, the blurred bodies were still visible, especially when a tracking error happened. As a result, users might have had difficulty focusing on their virtual avatars' appearance.

To more precisely examine the Proteus effect in the 3PP AR, the users' real body should be fully erased so that they can focus on their virtual avatar appearance. Therefore, I conducted a second study exploring the Proteus effect in 3PP AR when the users' real bodies are fully diminished. Furthermore, I compared a non-muscular and a muscular avatars to confirm the Proteus effect from an avatar's musculature as some VR studies showed the effect [33, 39]. The main hypothesis is that the Proteus effect of a muscular avatar would be induced in the 3PP AR environment. The study results found that the 3PP AR with real body removal induces the Proteus effect when using a muscular avatar in a physical performance task.

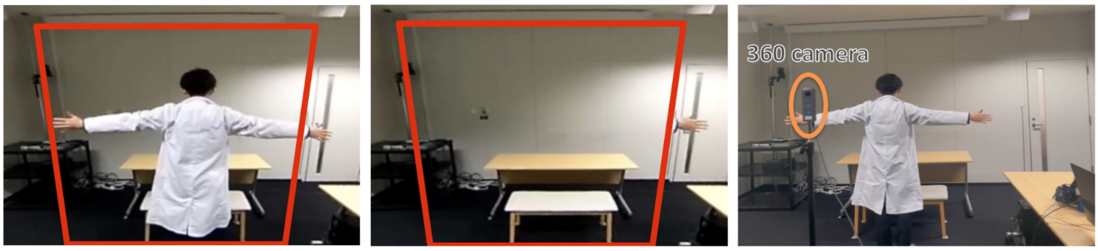

Figure 4.1: The user body is visible in a 360 panorama image before filling a certain area which was surrounded by red line (left). On the other hand, the user body is invisible in the certain area after filling this area (center) while the user is in this area in the real environment (right).

4.2 Methodology

4.2.1 Diminishing the user's body

The prototype system enables users to observe their avatars from the 3PP in AR. Before a user embody in the avatar, I run the system and render 360 images. Then, I stopped a certain area of the 360 image. As a result, the user can not see their real body as long as the body is within the area where rendering was stopped. Figure 4.1 illustrates how to diminish a user's real body. The size of the static image was set beforehand to cover the user's real body. In this way the user can no longer see their real body in the 360 video view.

4.2.2 Overlaying avatar body

A 3D virtual avatar was overlaid on the 360 image to represent the user (see Fig. 5.2). Both a non-muscular avatar and a muscular virtual avatar were created using Daz3D[*]. The virtual avatar had a dumbbell in its hand, and the position of dumbbell was synchronized with the HTC Vive tracker mounted on the real dumbbell was tracked. I employed Final IK[†] for an animation effect to hold up-down the dumbbell. As a result, the user could see that the avatar moved in sync

[*]https://www.daz3d.com/
[†]https://assetstore.unity.com/packages/tools/animation/final-ik-14290

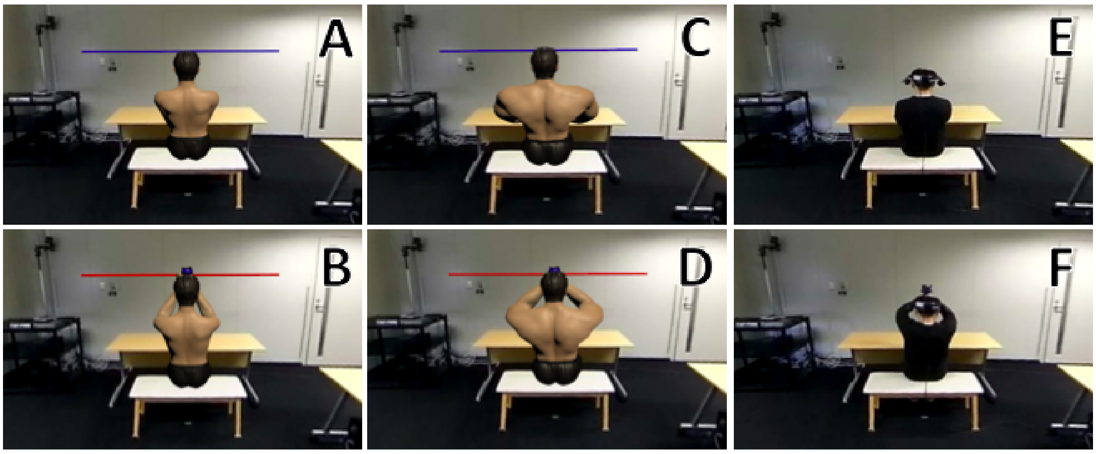

Figure 4.2: The user's view with the non-muscular avatar while holding the dumbbell on the table (A) and lifting the dumbbell up (B), the user's view with the muscular avatar while holding the dumbbell on the table (C), and lifting the dumbbell up (D), the original view while holding the dumbbell on the table (E), and while holding the dumbbell up (F).

with their real arm movement. Additionally, the user's head movement was also tracked by a six degree of freedom tracker in the HTC Vive Pro HMD, thus the avatar's head moved synchronously with the user's real head movement.

For the experiment task, the user was asked to lift up and down the virtual dumbbell while sitting in the real world. This was because the captured body parts were limited to the arms and the head. So, the user was not allowed to move their real legs because the virtual avatar's legs does not match the user's real leg movement.

4.2.3 Implementation details

In the prototype system I used real-time 360 panorama images captured by a Ricoh Theta V camera positioned at 1.5 meters above the ground and about 2 meters behind the user (see Fig. 4.3). The user wore the HTC VIVE Pro head-mounted display (HMD) to see the AR environment. The experimental task involved lifting dumbbell weights, and so I mounted an HTC VIVE Tracker on a 5kg dumbbell to capture the position of the dumbbell. In the experiment task,

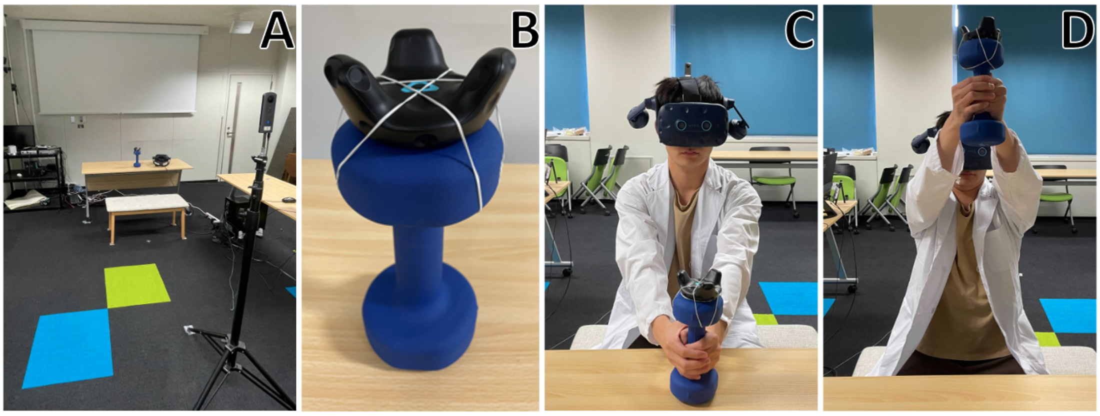

Figure 4.3: The 360 panorama camera was placed behind the bench where the user sat and conducted the task (A), and I prepared the dumbbell on which the tracker was mounted (B). The user held the dumbbell on the table (C) and lifted it up (D).

the user lifted up the dumbbell while viewing their virtual avatar embodiment. In addition, the system used a real-time 360 video that was rendered using the Unity game engine (2019.3.11), running on a PC with Windows 10.

4.3 User study

4.3.1 Experimental design

The study followed a within-subjects design to figure out the effect of a muscular avatar body appearance by comparing with a non-muscular avatar. The following conditions were compared and counterbalanced with a Latin-square design.

- **Condition NA**: The user was embodied as a non-muscular body avatar in the 3PP.
- **Condition MA**: The user was embodied as a muscular body avatar in the 3PP.

4.3.2 Experimental measurements

SoE self-report

I used the SoE self-report survey to examine the SoE toward an avatar, the same as used in study #1. The survey questions were taken from Peck's SoE questionnaire [**?**], and are listed in **??**. Previous studies have shown that when an avatar appearance is similar to the user, they feel a stronger SoE [41]. Hence, the experiment hypothesis is:

H1 Compared to the muscular avatar, participants would feel a higher sense of embodiment (SoE) for the non-muscular avatar.

The number of lift the dumbbell

To examine the Proteus effect of a muscular avatar, I measured how many times participants could lift the dumbbell up and down. If they receive the Proteus effect from a muscular avatar, their physical performance might be improved, and they would lift the dumbbell more times. As a result, I hypothesised:

H2 Compared to a non-muscular virtual avatar, using a muscular virtual avatar would improve the user's physical performance while controlling the avatar.

Post-hoc hand grip strength

In addition to the immediate Proteus effect while being embodied, I also evaluated the prolonged Proteus effect after the avatar embodiment. If the Proteus effect of a muscular avatar remains, the user's physical performance should still be enhanced after the avatar embodiment. Hence, I hypothesized:

H3 Compared to using a non-muscular avatar, using a muscular avatar would improve the user's physical performance after controlling the avatar.

4.3.3 Procedure

After the participants filled in a pre-questionnaire about their body height and VR experience, they sat on the center of a bench and put on the HMD. During the AR experiment, the participants could see their avatar in a real-time 360 video of the real environment with their real body removed. They were instructed to lift up and down a 5 kg dumbbell (Umi Dumbbell) five times as a practice. Then, the participants were asked to lift the dumbbell up and down every five seconds as many times as possible.

After the participants decided to stop lifting the dumbbell, they took off the HMD and took a break for two minutes. I determined the break time length through a pilot study. After the two minutes break, the participants left the bench and their hand-grip strength was measured with a dynamometer (N-FORCE HG-251).

Next, the participants answered a questionnaire evaluating their SoE with the questionnaire [**?**], which consisted of 16 questions (7-point Likert scale, 1 = strongly disagree, 7 = strongly agree) with four sub-scales (appearance, response, ownership, and multi-sensory) (see **??** for the full questionnaire). After that, the participants repeated the task using the other avatar and filled in the questionnaire about the SoE again.

4.3.4 Participants

A total of 16 male participants aged between 22 and 34 years old (M = 23.69, SD = 2.87) participated in this study. All but one of the participants were right-handed, and all participants had experienced VR at least once.

4.4 Results

4.4.1 SoE self-report

The SoE was analyzed using a Wilcoxon signed-rank test to verify **H1**. There was no significant differences between the two body shape conditions in terms of SoE (W=52, p=0.43) (see Fig. 4.4 (left)).

4.4.2 The number of up-down lifts of the dumbbell

The number of up-down lifts of the dumbbell was analyzed using a Wilcoxon signed-rank test to verify **H2**. As expected, the average number of up and down

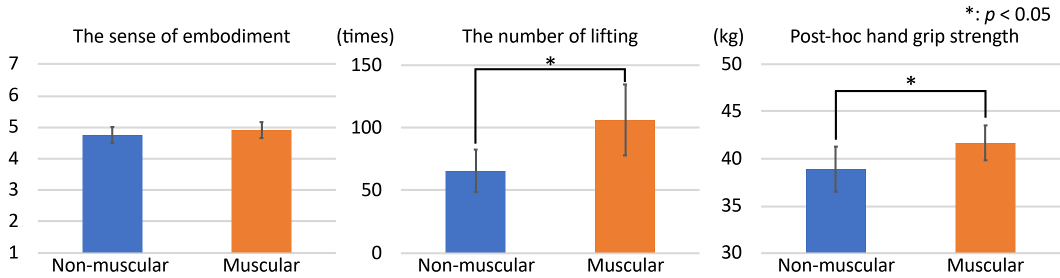

Figure 4.4: The average sense of embodiment measured on a Likert scale from 1 to 7 (left), the average number of up-down lifts of the dumbbell (middle), and the average post-hoc hand grip strength (kg) (right).

lifts of the dumbbell was significantly larger in the muscular avatar condition than in the non-muscular avatar condition (W=25, p=0.02) (see Fig. 4.4 (middle)). Hence the Proteus effect was induced while being embodied in the muscular avatar.

4.4.3 Post-hoc hand-grip strength

The post-hoc hand-grip strength was analyzed using a paired t-test to verify **H3**. As expected, the post-hoc hand grip strength was significantly higher after being embodied in the muscular avatar than in the non-muscular avatar ($t(15) = -2.84$, p=0.01) (see Fig. 4.4 (right)). Therefore, the Proteus effect was maintained even after the end of the embodiment phase in terms of hand-grip strength.

4.5 Discussion

The results did not support **H1** as there was not a significant difference between a non-muscular and a muscular avatars in terms of the average SoE. This finding indicates that visuo-motor synchronization between the user and the avatar might help the user feel the same SoE toward a muscular avatar as a non-muscular avatar. **H2** and **H3** were supported, thus it was shown that the muscular avatar from the 3PP in AR improves user physical performance during and after controlling the avatar. For example participant P7 said that "I felt controlling the muscular avatar so that lifting the dumbbell was not hard" and participant P16 commented that " I felt like I became muscular while being embodied in a muscular avatar, then I was motivated to overcome the limitations".

However, there were some limitations in this work. First, the participants were instructed to lift the dumbbell up and down as many times as possible in each task. Thus, the duration of the AR experience and avatar embodiment was different for each participant. This also could have produced a different degree of tiredness in each participant which could have affected the grip strength test.

For example participant P7 commented that "deciding when I stop lifting the dumbbell was difficult". Therefore, there is a possibility that some participants quit lifting before they reached their physical limits.

Participants also might have gotten tired after the first condition which impacted their performance in the second condition. For example, participant P4 suggested "lifting the dumbbell to my limit twice was so hard".

There were also limitations in the subject pool chosen. No female participated in this experiment, thus the Proteus effect from the 3PP in AR for female users remains unanswered. The subjects were also aged between 22 and 34 years old, and so there was no insight gained about the Proteus effect for older people.

4.6 Conclusion

In this study, a prototype 3PP AR system was developed using a synchronized avatar integrated into a real-time 360 video where their real body is completely removed. Using the prototype system, I validated the SoE using the questionnaire which was used in the study #1. I examined the Proteus effect of a muscular avatar by measuring the number of times a user could lift up and down a dumbbell and their post-hoc hand grip strength respectively. The results showed that the user could feel the same SoE when embodied in a muscular avatar as when using a non-muscular avatar in the 3PP AR. Embodied in a muscular avatar from the 3PP in AR improved the user's physical performance both during and after embodiment.

Unlike study #1 where the users body was blurred but still visible, the user's body was completely erased by using a pre-captured image in the real-time 360 video. As a result, the Proteus effect was supported in the 3PP AR while it was not supported in the study #1. This could be because that completely erased user body helped focus on the avatar appearance, thus the user could feel the Proteus effect more than when they are distracted by their own bodies. This result would motivate research on inpainting technology to improve the AR embodiment experience.

In this study, I only hired male participants due to recruiting limitations. So in the future, a follow-up study needs to be conducted with female users and 1PP. Furthermore, other experimental measures should be used to examine user physical performance. As mentioned in 6, counting the number of times to lift the dumbbell relied on participants' decisions. Therefore, the evaluation of the Proteus effect was subjective, and objective methods would be required to examine the Proteus effect more precisely. For instance, physiological sensor such as sEMG could be used to measure the physical performance.

The main finding of the study #2 is the immediate and prolonged Proteus effect in the 3PP AR. This results encouraged me to come up a method to enhance the Proteus effect in AR. As a result, I studied visual transition in the following study.

5 Study #3: the impact of visual transition under fully erased user body

5.1 Overview

Study #2 showed that embodying a muscular-shaped avatar improves the user's physical performance. In order to benefit from these diverse enhancements in the real world, one could imagine using an AR embodiment system enabling the switch between these different avatars at will.

However, abruptly switching avatars may disrupt the continuity of the embodiment experience and impact the SoE. In addition to this, it is not guaranteed that the new avatar will be able to evoke an as strong SoE. Indeed, previous work has often shown that this ability varies highly depending on users and avatars [10]. Although how the avatar's appearance impacts the SoE is still not completely clear, more and more work suggests that this might be linked to the avatar's appearance as those that differ a lot from their users tended to produce weaker embodiment illusions [11, 24, 51, 86].

In this chapter, I propose to address these issues by implementing visual transitions at the moment of the avatar's change. I designed a system letting users smoothly transition from one avatar to another through visual effects. Our approach relies on AR to display the avatar in the third-person perspective so that the user can fully witness the transformation within a real-world context. In order to launch the transformation, I explored two approaches: one where the user is in charge of triggering the transition through a specific action (referred to as an "active" transition), and one where the user sees the avatar change automatically ("passive" transition).

To evaluate the effects of these transitions on how the SoE evolves when switching avatars, I ran a user study where the user goes from an avatar with a body similar to their own to a highly muscular-shaped avatar. I hypothesized that applying a smooth transition through our system would allow maintaining the SoE at the level of the regular-shaped avatar when switching to the muscular-shaped avatar. As a consequence, I imagined that the resulting Proteus effects could be enhanced compared to a control condition where visual transitions were not applied.

Our results suggest that applying active visual transitions, compared to passive ones, can benefit the stability of the SoE when switching avatars. They also partially support that Proteus effects induced by a muscular-shaped avatar were enhanced thanks to this transition. In particular, our contributions can be summarized as follows:

- I present a novel AR system providing an end-to-end embodiment experience starting from being in one's real body to embodying different types of avatars in the third-person perspective before returning to one's real body.
- I present a user study investigating the effects of applying visual transitions when changing one's avatar on the evolution of the SoE and Proteus effects.

5.2 Transition system

In the study, I aimed to address our research question of how visual transitions impact a user's SoE and the Proteus effect through end-to-end avatar embodiment experience in AR. To do that, I created a 3PP AR embodiment system that combines several types of transitions at different stages of the embodiment. First, as suddenly seeing the world from a 3PP can be unsettling, I implemented a visual transition allowing users to switch perspectives with a fast camera translation. I refer to these as perspective transitions. Second, instead of directly displaying the user's avatar in front of them, I propose to let them first see the 3D mesh of their real bodies before gradually fading into the avatar (regular-shaped). Next,

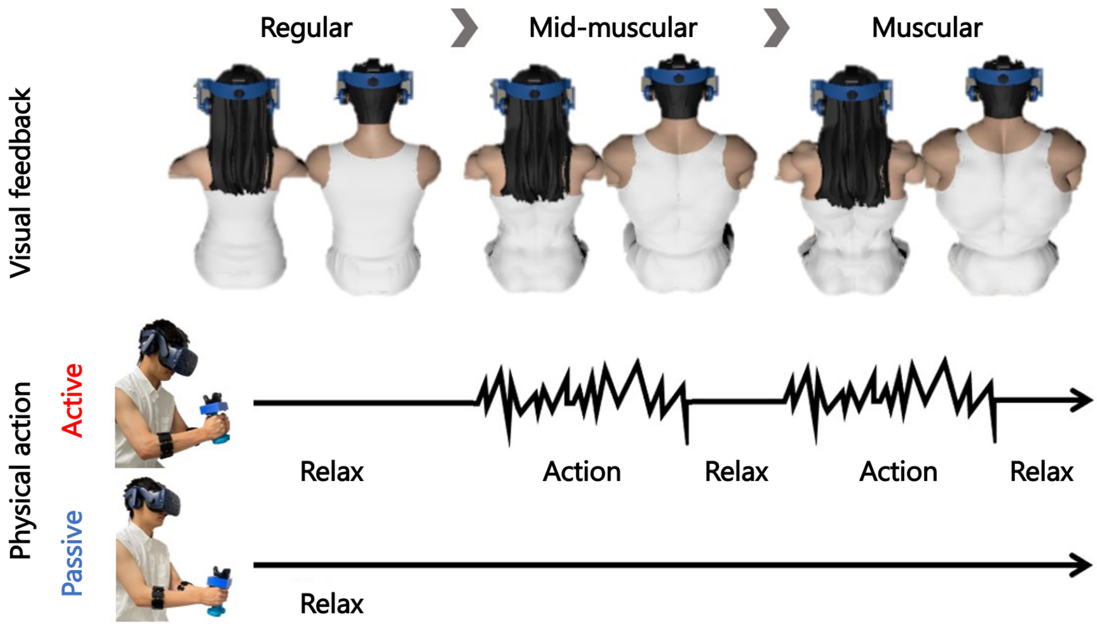

Figure 5.1: Illustration of the visual transitions I implemented to change the appearance of one's avatar. *Visual feedback:* during the transition, the user can see their avatar transforming progressively from a regular-shaped body to a muscular-shaped one. *Physical action:* to be experienced, this transition can be either triggered by the user through physical action (active) or launched automatically (passive).

the user may switch from this avatar to another one (muscular-shaped) through a visual transition that gains an avatar's muscle. Finally, users are able to return to their real bodies through the same process the other way around. Figure 5.5 shows a detailed timeline of these different phases.

To trigger the transitions between two virtual avatars, I explored two approaches :

- **Active transitions**: in this approach, the user may trigger the change of appearance by performing a physical action (squeezing dumbbells). They control the progress of the transformation thanks to visual feedback synchronized with their action.
- **Passive transitions**: the change of avatar is performed automatically by the system and the user witnesses its progress without controlling it.

Both of these approaches can be interesting depending on the targeted embodiment experience or context. For instance, when designing a game that involves physical exertion, one could desire to make the user's avatar appearance evolve based on the physical activity of the player (e.g., in a fitness game like Ring Fit Adventure on Nintendo Switch, it would become more and more athletic). In other situations such as when simply exploring different avatar appearances for personalization done later, passive transitions could be preferred.

5.2.1 Implementation details

The system I present was implemented using Unity (2019.3.11f) and ran on a Windows 10 computer equipped with an NVIDIA GeForce GTX 1080 Ti graphics card. I used an HTC VIVE Pro Eye [*] head-mounted display (HMD) to display the avatars and visual transitions.

Avatars. The 3D mesh of the user's real body is captured with the help of an RGB-D camera (Microsoft Azure Kinect [†]) and overlaid onto the 360 videos. The Azure Kinect DK was positioned right next to the RICOH THETA V. The 3D models I used to test visual transitions between avatars were regular-shaped female and male avatars and their muscular-shaped versions (Fig. 5.2). They were designed based on the average Japanese body size using Daz3D [‡], from the characters "Genesis 8 Female" and "Genesis 8 Male".

Arm tracking. To track the user's arm movements, users were asked to hold onto an object equipped with an HTC VIVE tracker (3.0) [§]). This object joined two 1-kg weights together with a plastic piece so that the user had to grasp both weights to lift it (see Fig. 5.4).

Active transition trigger. Users were required to wear Myo armbands (Thalmic Labs), a wearable electronic device that detects a user's hand gestures through

[*] https://www.vive.com/nz/product/vive-pro-eye/overview/
[†] https://azure.microsoft.com/en-us/products/kinect-dk/
[‡] https://www.daz3d.com/
[§] https://www.vive.com/nz/accessory/tracker3/

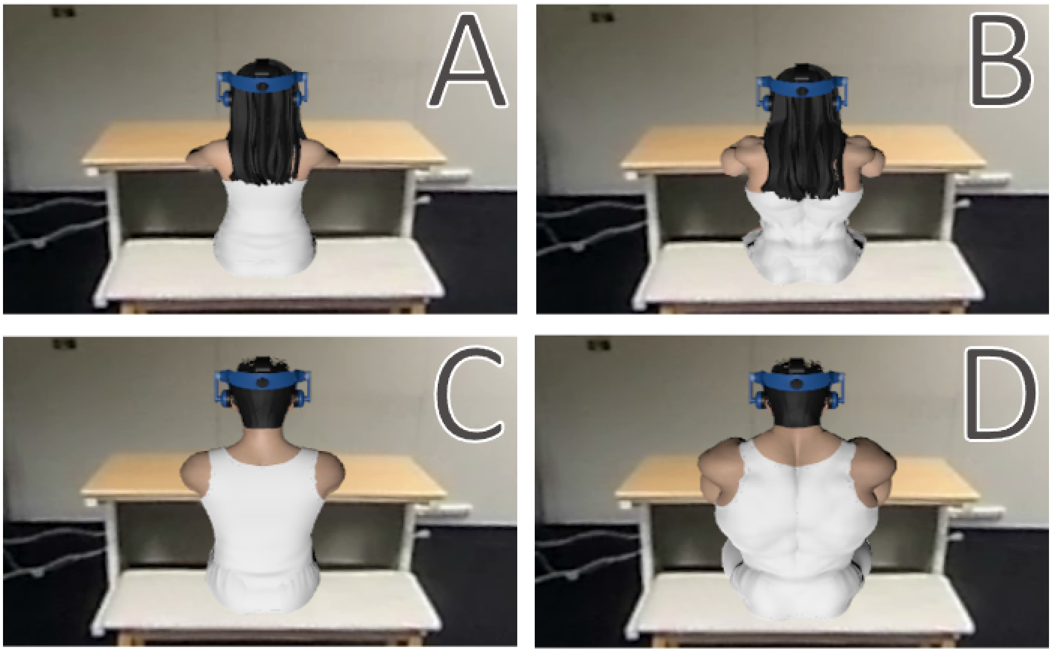

Figure 5.2: The upper left (A) and right (B) are female regular- and muscular-shaped avatars, respectively. The bottom left (C) and right (D) are male regular- and muscular-shaped avatars, respectively.

sEMG [2, 45, 47, 80]. These sensors were worn on their forearms and captured real-time sEMG values from the extensor carpi ulnaris (ECU) muscle at 50 Hz

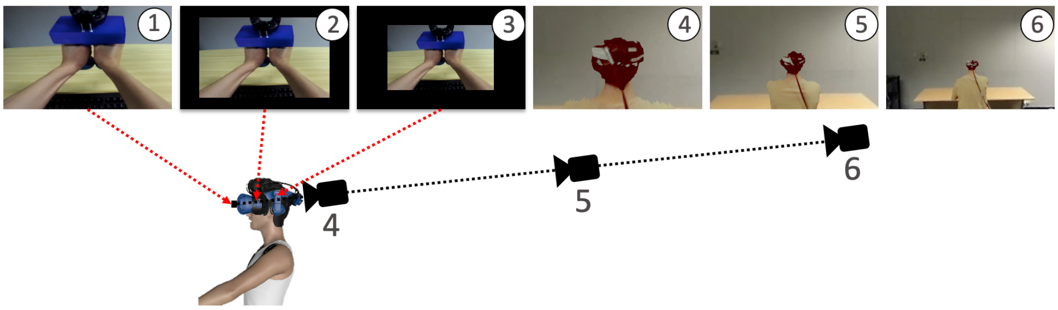

Figure 5.3: The path of a virtual camera looking at the avatar and the corresponding view. The user is in the 1PP at the beginning (1 to 3), then in the 3PP (4 to 6) in transitions from the 1PP to the 3PP. The user gets through the reverse path in transitions from the 3PP to the 1PP.

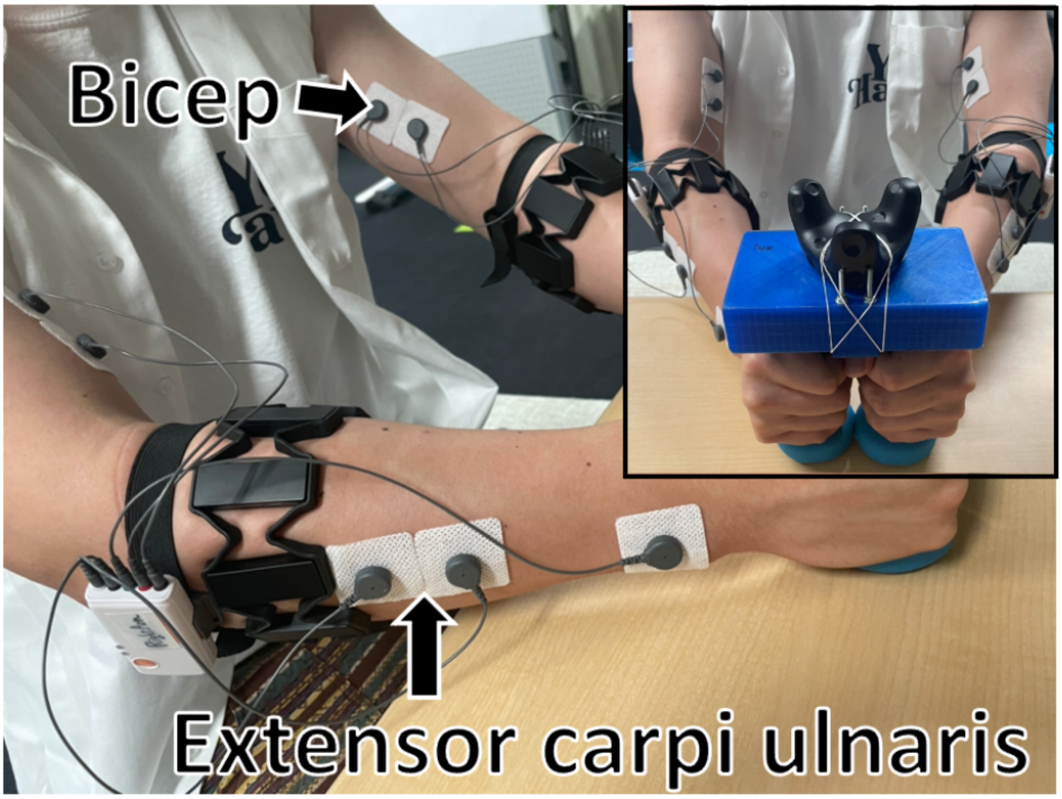

Figure 5.4: The sEMG sensor electrodes were attached to the user's ECUs and biceps. Participants held the dumbbells with both hands. The tracker was mounted on the dumbbell pair to track the user's arm movement.

through Bluetooth. The user's squeezing action could then be detected based on the real-time sEMG values.

In addition to these Myo armbands, I used a Shimmer3 EMG unit ¶ for offline analysis of the sEMG magnitude at 512 Hz through Bluetooth.

AR visualisation. While wearing the HMD, users could observe the real world around them in the 1PP through a stereo camera attached to the HMD (ZED Mini ‖). The 3PP, on the other hand, was provided thanks to real-time 360 videos captured by a RICOH THETA V camera ** positioned 1.5 meters above

¶https://shimmersensing.com/product/shimmer3-emg-unit/

‖https://www.stereolabs.com/zed-mini/

**https://theta360.com/en/about/theta/v.html the ground and about two meters behind users.

5.2.2 Perspective transitions

The transition between the 1PP and the 3PP was achieved by moving the virtual camera from the user's 3D mesh head position to the centre of a virtual sphere onto which the real-time 360 videos of the environment were rendered (see Fig. 5.3).

When going from the 1PP to the 3PP, the real-time video of the stereo camera zooms out first. Then, the virtual camera moves back to the centre of the virtual sphere rendering the real-time 360 videos. When going from the 3PP to the 1PP, the virtual camera moves forward from the centre of the virtual sphere to the avatar's head first. Then, the real-time video of the stereo camera zooms in. Each perspective transition takes 10 seconds to be achieved.

5.2.3 Transitions from a user mesh to a regular-shaped avatar

Once in the 3PP, users can view the 3D mesh of their real bodies and see it change into a regular-shaped avatar through the 360 real-time videos. The 3D mesh is generated from depth data retrieved with the Azure Kinect development kit, as done in previous work [49]. The transition to the regular-shaped avatar is done by gradually lowering the number of vertices in the mesh to display a "dissolving" effect. At the same time, the regular-shaped avatar is gradually made visible, from fully transparent to fully opaque.

Lastly, to avoid the user's actual body image being caught by the 360 camera, a portion of the 360 images was substituted with a static image taken beforehand. Regarding the 3D mesh, using depth-based background subtraction, only the user's actual body is rendered on the 360 video.

5.2.4 Transitions from a regular- to a muscular-shaped avatar

I change the regular-shaped avatars into muscular ones by adjusting morph targets[††]. The avatar's body weights (ABW) of the female and male regular-shaped avatars are estimated using the formula of Buckley et al. (2010) [4]. For the regular-shaped avatar, the height and weight are 160.1 cm and 53.8 kg for a female, and 173.5 cm and 71.6 kg for a male. For the muscular-shaped avatar, the weight is 77.4 kg for a female and 97.8 kg for a male (the height remained the same). As a result, the muscular-shaped avatars are approximately 40% heavier than the regular-shaped ones.

Active transitions. To make their avatar become muscular, users have to repeat the physical actions of firmly squeezing the dumbbell pair and then relaxing while still holding it. While doing this, the avatar appears to become more muscular every time they squeeze. This is referred to as the "visual feedback" of the transformation. The avatar's muscularity increased by 20% when the sEMG values exceeded the 1.5 times users' average values, which are measured beforehand. After the users flex their muscles up to five times, the avatar's body reaches its maximum muscle mass.

Passive transitions. In this case, the users could see their avatar's body becoming more muscular without any action on their part. Their avatar gained 20% of muscle mass every five seconds and reached its maximum muscle mass after 30 seconds, which was about the same time required by an active transition.

5.3 User Study

Through this user study, I investigated the effects of using visual transitions when changing from a regular- to a muscular-shaped avatar on the SoE and potential Proteus effects.

[††] https://www.daz3d.com/massive-morphs-for-genesis-8-male, https://www.daz3d.com/massive-morphs-for-genesis-8-female-s

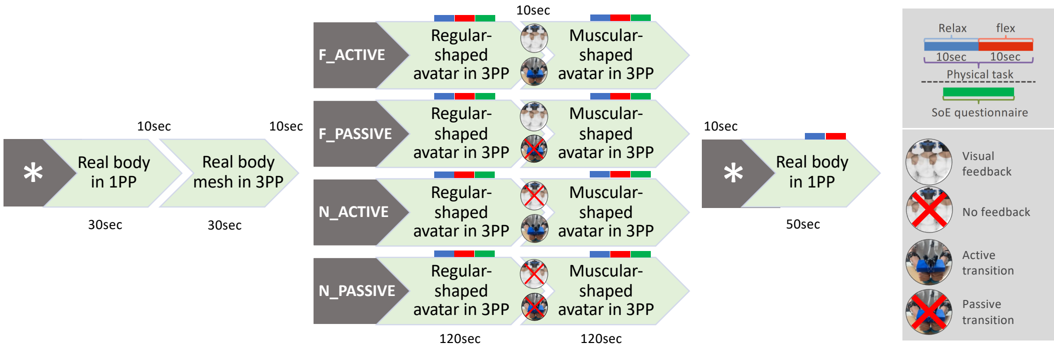

Figure 5.5: The experimental procedure. This figure details the timing of tasks in each condition, including the users' viewpoint, self-report, and sEMG magnitude measurements.

5.3.1 Experimental design

To check the effects of such transitions on the SoE, I needed at least 2 conditions: a control condition without any transition, and a test condition where a visual transition was applied. As I were interested in comparing the effects of both active and passive transitions, I decided to split our test condition into two separate ones ("physical action" factor).

Additionally, it appeared to us that I would need to investigate how visual feedback, which allows a user to observe their regular-shaped avatar appearance gradually change to the muscular-shaped one, impacts the difference in the SoE and the intensity of the Proteus effect. Therefore, it seemed that the presence/absence of visual feedback during an avatar's appearance change could be one variable to include in our experiment ("visual feedback" factor).

Consequently, I performed a cross-evaluation of 2 × 2 (visual feedback *vs.* no feedback × active transition *vs.* passive transition). To do so, I designed four conditions applied in a within-subjects design:

- **Condition F_ACTIVE**: with visual feedback and an active transition. The user observes their avatar's appearance progressively become more muscular through their actions.
- **Condition F_PASSIVE**: with visual feedback and a passive transition. The appearance of the user's avatar automatically gains muscle without any

physical effort from the user.

- **Condition N_ACTIVE**: with no visual feedback and an active transition. The user repeats the physical action, yet their view is masked by a black screen until the end of the transformation.
- **Condition N_PASSIVE**: with no visual feedback and a passive transition. The user's view is masked by a black screen and they wait for 30 seconds without doing anything to see their new avatar.

All subjects experienced all of four conditions, but the order of the conditions was counterbalanced with a Latin-squares design to remove any order effects.

5.3.2 Measures

I used self-report questionnaires to gauge the SoE across avatars and conditions. I also evaluated potential Proteus effects by measuring the magnitude of the sEMG.

Sense of embodiment

The participants were asked to answer six questions on a 7-point Likert scale, relating to different components of the SoE (see Table 5.1). The questions were taken from a study investigating 3PP avatar embodiment and allowed computing scores for the sense of agency, body ownership, and self-location [15]. To measure the difference in the SoE between regular- and muscular-shaped avatars, I needed the participants to answer the SoE questionnaire while embodying each avatar (i.e., two times per condition, and eight times in total through one experiment). To do so, I displayed the questionnaire virtually in front of them and used the eye-tracking system embedded in the HMD to allow them to answer the questions without removing their hands from the dumbells (see Fig. 5.6). Participants could choose the number best representing their experience using a virtual ray controlled by their eye-gaze. To confirm their answer, they had to keep looking at the same number for 3.0 seconds.

As mentioned in Section **??**, previous research suggests that avatars whose appearance differs a lot from the user tend to induce a weaker SoE [11, 24, 51].

Table 5.1: The items of SoE self-report based on Debarba et al. (2017) [15]. Answers were on a 7-point Likert scale ranging from 1 (completely disagree) to 7 (completely agree).

Item number	Category	Question
Q1	Agency	It felt like I was in control of the body I was seeing.
Q2	Agency	Whenever I moved my body I expected the virtual body to move in the same way.
Q3	Ownership	I felt as if I was looking to my own body.
Q4	Ownership	It felt that virtual body was my own body.
Q5	Self-location	It felt as if my body was located where I saw the virtual body to be.
Q6	Self-location	It seemed as if I were sensing the movement of my body in the location where the virtual body moved.

Therefore, I expected that in our control conditions, the SoE would naturally decrease from the moment the user embodied the regular-shaped to the one they embodied the muscular-shaped avatar. Therefore, to verify whether applying transitions could help maintain a constant (or less diverging) SoE, I computed the differences in SoE scores before and after changing the avatar's appearance for each condition. I subtracted each SoE component score of the regular-shaped avatar from the corresponding score of the muscular one using the following equation:

$$\Delta Q_N = score(Q_N)_{muscular} - score(Q_N)_{regular} \quad (5.1)$$

$$\Delta_{Agency} = (\Delta Q_1 + \Delta Q_2)/2 \quad (5.2)$$

$$\Delta_{Ownership} = (\Delta Q_3 + \Delta Q_4)/2 \quad (5.3)$$

$$\Delta_{Self\_Location} = (\Delta Q_5 + \Delta Q_6)/2 \quad (5.4)$$

In the above, $N = 1, 2, ..., 6$, $score(Q_N)_{muscular}$ is the answer of the Nth question when embodying the muscular-shaped avatar, and $score(Q_N)_{regular}$ is the same for the regular-shaped avatar.

Participants are most likely to answer the second SoE questionnaire relative to the first one. Therefore, I expected such a measure would be more sensitive

Figure 5.6: The users view while answering the question with eye-gaze interaction.

to how the SoE varied based on the transition applied. Additionally, I had the intuition that the active transitions would provide a higher sense of agency than the passive condition (as the user controls the transformation), and therefore generate a superior embodiment. Our hypotheses can be summarized as follows:

H1.1 F_ACTIVE and F_PASSIVE will provide a higher Δ_{Agency}, $\Delta_{Ownership}$, and $\Delta_{Self\_Location}$ than N_ACTIVE and N_PASSIVE.

H1.2 F_ACTIVE will provide a higher increase in Δ_{Agency}, $\Delta_{Ownership}$, and $\Delta_{Self\_Location}$ than F_PASSIVE.

sEMG magnitude

As suggested by previous work [33, 39], I assumed that embodying a muscular-shaped avatar would lead to Proteus effects consisting of users applying more strength during a physical task. Therefore, I included a physical task in all conditions to monitor the presence of such effects. The task consisted in grabbing the dumbbell pair as tightly as possible for 10 seconds. In each condition, participants were asked to perform this task three times: during the embodiment of the regular-shaped avatar, of the muscular-shaped avatar, and after the embodiment.

I captured the sEMG magnitude during the tasks over Bluetooth (512 Hz) thanks to the sensors attached to the ECU and biceps muscles (see Fig. 5.4). This allowed us to evaluate the user's physical performance across avatars and conditions. After the experiment, the acquired sEMG values were band-pass filtered from 25 to 200 Hz and the root means square (RMS) was calculated using a sliding window of 100 milliseconds following the methodology used in the study of Perusquía-Hernández et al. (2021) [56].

The period when participants contracted their muscles during the physical task was analyzed using nine-second epochs. The first second after the beep signal indicating when to start squeezing was removed to exclude the transient section of the sEMG for the duration of the beep (1 second). The timing of the beep signal and the Shimmer3 EMG data capture were synchronized using a hardware trigger sent via the serial port from Unity. The average RMS activity was calculated per epoch, and the ratio of the average RMS activity was calculated as the normalized sEMG magnitude. The normalized RMS activity when embodied in a muscular-shaped avatar (EVR_{during}), and after embodiment (EVR_{after}) was considered using the following equation:

$$EVR_{during} = \frac{EMG_{muscular}}{EMG_{regular}} \tag{5.5}$$

$$EVR_{after} = \frac{EMG_{after\_embodiment}}{EMG_{regular}} \tag{5.6}$$

where $EMG_{regular}$ is the average RMS during embodiment in the regular-shaped avatar, $EMG_{muscular}$ is the average RMS during embodiment in the muscular-shaped avatar, and $EMG_{after\_embodiment}$ is the average RMS after embodiment.

The average of the dominant hand's ECU's EVR_{during} and the bicep's EVR_{during} was calculated as the dominant hand's EVR_{during}. The average of the non-dominant hand's ECU's EVR_{during} and the bicep's EVR_{during} was used as the non-dominant hand's EVR_{during}. In the same way, the dominant hand's EVR_{after} and the non-dominant hand's EVR_{after} were calculated using the EVR_{after} of dominant hand's ECU and bicep, and the EVR_{after} of non-dominant hand's ECU and bicep respectively.

I hypothesized that using transitions with visual feedback from a regular- to a muscular-shaped avatar would enhance the potential Proteus effect as the SoE is assumed to be increased. Thus, the user's physical performance would be higher during the embodiment of the muscular-shaped avatar compared to the control conditions at equivalent moments. I also hypothesized that applying transitions with visual feedback would allow prolonging the Proteus effect after embodiment ended, as shown possible in prior research [33,39]. Lastly, I expected that visual feedback synchronized with an active transition would boost the impact of the Proteus effect more than with a passive transition. The hypotheses about the sEMG magnitude measures can be summarized as follows:

H2.1 F_ACTIVE and F_PASSIVE will result in increases of EVR_{during} for both dominant and non-dominant hands.

H2.2 F_ACTIVE will further enhance both dominant and non-dominant hands' EVR_{during} than F_PASSIVE.

H3.1 F_ACTIVE and F_PASSIVE will result in increases of EVR_{after} for both dominant and non-dominant hands.

H3.2 F_ACTIVE will further enhance both dominant and non-dominant hands' EVR_{after} than F_PASSIVE.

5.3.3 Procedure

The experiment started by collecting the informed consent form signed by the participant. They were then asked to complete a demographics questionnaire retrieving data on their body weight and previous VR experience. To match the avatar's clothing, the participants were then instructed to wear a white sleeveless shirt. Participants were then equipped with a Shimmer3 EMG sensor and a Myo sensor on each of their arms (see Fig. 5.4). After that, the experimenter gave a brief introduction to the goal of the experiment and their task. I highlighted that the participants could withdraw from the experiment at any time without penalty or losing their monetary reward.

Next, I collected baseline Myo sensors values to allow comparing the sEMG data while embodying the avatar in the analysis. To do so, participants were instructed to sit in the centre of the bench and hold the dumbbell pair for 10 seconds twice. In the first measurement, participants were asked to relax while holding the dumbbells for 30 seconds, while in the second measurement, they were asked to hold the dumbbells as tightly as possible for 30 seconds. After that, eye-tracking calibration was carried out and participants took a break for two minutes. Then, the experimenter launched the experimental software on the HMD in video see-through 1PP mode (i.e., using the Zed Mini cameras) and asked the participant to wear it. At this point, the user could see an AR view of the experimental room. They were asked to look around while passively holding the dumbbell pair.

The main part of the experiment started from here, following the steps described in Fig. 5.3. After 30 seconds in the 1PP, the perspective transition was triggered to switch to the 3PP. Participants could then observe their real bodies as a 3D mesh in front of them. They were asked to look around again and catch the experimenter waving at them on the right to further show that the scene was live.

After 30 seconds in the 3PP, the real-body mesh dissolved to transition to a generic regular-shaped avatar matching the participant's gender. The participant was instructed to look around and lift the dumbbell pair up and down twice. This allowed the participants to confirm that their avatars moved synchronously with them. Then, the participant went through a physical task during which they had to contract their muscles upon hearing three "beep" sounds, separated by 10 seconds. The first beep sound indicated that the participant should get ready while relaxing their arms. After the second beep sound, they had to hold the dumbbells as tightly as possible until the third beep sound occurred. Following this, participants could relax again and were asked to answer the SoE questionnaire displayed in front of them using eye-gaze interaction (see Fig. 5.5).

Depending on the condition, the participant then saw (or not) their regular-shaped avatar morph into a muscular one, either through an active transition or a passive one as described in Section 5.2.4. Participants performed the physical task again and completed the SoE questionnaire once more while observing their muscular-shaped avatar. After that, participants returned to the 1PP through a perspective transition, looked around, and went through the same physical task again. They could then remove the HMD and had a two-minute break to rest before starting all over for the next condition. Once all conditions were finished, participants provided feedback and received their reward.

5.3.4 Participants

A total of 32 volunteers (16 females) were recruited via university mailing lists. Participants aged between 21 and 35 years old ($M = 24.94$, $SD = 3.40$) took part in this study. All but three of the participants were right-handed, and 22 participants had experienced VR at least once. Each participant was paid 1000 JPY as a reward.

5.4 Results

Our analysis covered the parameters summarized in Table 5.2. Shapiro-Wilk tests showed that the (subtracted) SoE scores and sEMG data deviated from the normal distribution. Therefore, all of the parameters were analysed using Friedman tests followed by Wilcoxon signed-rank posthoc tests. A p-value less than 0.05 was considered statistically significant. Fig.5.7 provides descriptive information on these parameters.

Table 5.2: The list of parameters' definitions.

Parameter	Definition
Δ_{Agency}	How much sense of agency for avatar is maintained.
$\Delta_{Ownership}$	How much sense of ownership for avatar is maintained.
$\Delta_{Self\_Location}$	How much sense of self-location for avatar is maintained.
EVR_{during}	Normalized sEMG magnitude during embodiment.
EVR_{after}	Normalized sEMG magunitude after embodiment.

5.4.1 Assumption checks

I expected that the muscular-shaped avatar would induce a lower SoE based on previous studies [11, 24, 41, 51, 86]. To verify this, I compared its absolute SoE scores with that of the regular-shaped avatar in the control condition (N_PASSIVE). A Wilcoxon signed-rank test showed that the ownership score for a muscular-shaped avatar was indeed significantly lower than for a regular-shaped one ($p < .05$), while other scores were not significantly different (agency: $p = .50$, self-location: $p = .20$).

From previous work [33, 39], I also assumed that the muscular-shaped avatar would induce a Proteus effect characterized by users employing greater force during physical tasks. To check whether this was the case, I compared the absolute sEMG values collected during and after the embodiment of the muscular-shaped avatar to during the embodiment of the regular-shaped one. As a result, I found that the EMG magnitude of right bicep muscle was significantly higher when users embodied in the muscular-shaped avatar than during the embodiment toward the regular-shaped one ($p < .05$). Therefore, our assumption that the Proteus effect of the muscular-shaped avatar was partially supported. However, I could not find significant difference in the sEMG values between the regular avatar and after embodiment.

5.4.2 Sense of embodiment

The Friedman test revealed the existence of significant differences in Δ_{Agency} ($\chi^2(3, \text{n} = 32) = 29.27$, $p < .001$). A Wilcoxon signed-rank test showed that these significant differences were between F_ACTIVE and F_PASSIVE ($p < .05$) and F_ACTIVE and N_ACTIVE ($p < .001$), with F_ACTIVE providing a superior sense of agency. Differences were also found between N_ACTIVE and N_PASSIVE ($p < .001$), with N_PASSIVE providing a superior agency (see Fig. 5.7, left).

On the other hand, the Friedman test did not show significant differences in terms of body ownership and self-location (ownership: $p = .18$, self-location: $p = .14$).

These results show that the presence of visual feedback (allowing one to see

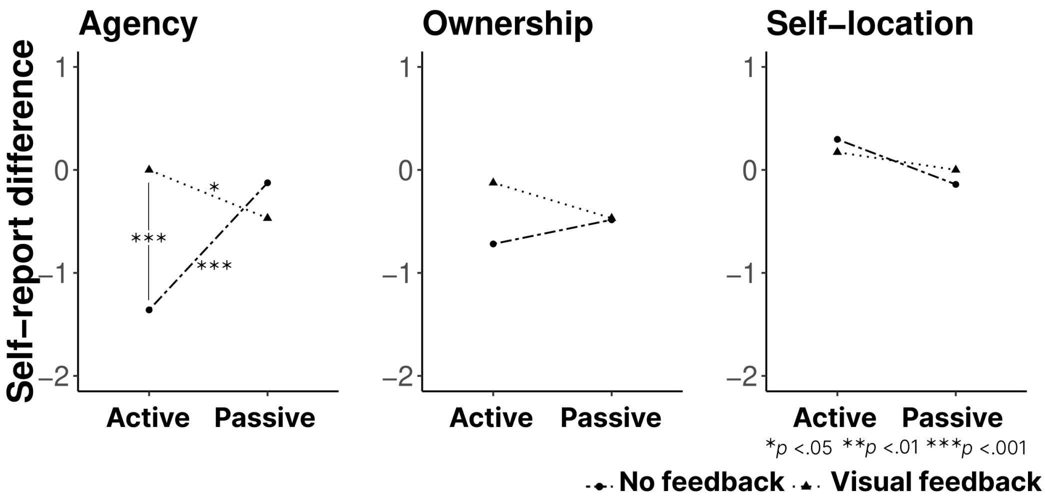

Figure 5.7: From left to right, the results of agency difference (Δ_{Agency}), the difference of ownership ($\Delta_{Ownership}$), and the difference of self-location ($\Delta_{Self\_Location}$). The result of (Δ_{Agency}) (Left) shows that visual feedback helped keep the sense of agency when an active transition was applied and an active transition with visual feedback retained the sense of agency more than a passive one with visual feedback.

the avatar progressively transform) has a positive impact on the maintaining of the sense of agency when an active transition is employed. This partially supports hypothesis **H1.1**. Compared to passive transitions with visual feedback, I also found that active transitions with visual feedback provided better help to maintain the sense of agency when switching from the regular- to the muscular-shaped avatar. This partially supports hypothesis **H1.2**.

5.4.3 sEMG magnitude with the muscular-shaped avatar

The Friedman test revealed significant differences between EVR_{during} of the non-dominant hand across conditions ($\chi^2(3, n = 32) = 9.15$, $p < .05$). The Wilcoxon signed-rank test showed that their origin was between F_ACTIVE and F_PASSIVE ($p < .01$) and F_ACTIVE and N_ACTIVE ($p < .05$), with F_ACTIVE providing superior sEMG values (see Fig. 5.8, top).

On the other hand, the Friedman test did not show significant differences in the EVR_{during} of the dominant hand ($p = .14$).

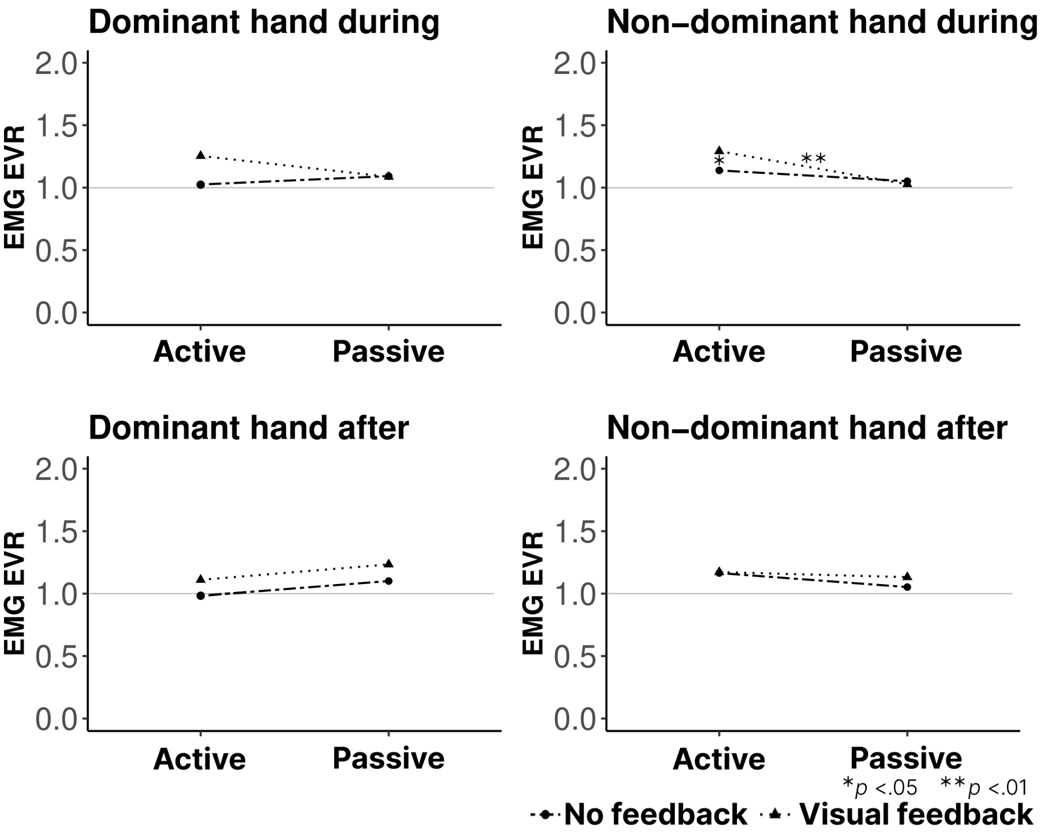

Figure 5.8: Upper plots depict the measurements during the embodiment. Lower plots depict the activity after the embodiment. Dominant and non-dominant hands' results are presented separately. The result of the non-dominant hand during (upper right) shows that visual feedback enhanced the sEMG magnitude and an active transition with visual feedback had a more positive impact on the sEMG during embodiment than a passive transition with visual feedback.

These results show that the usage of active transitions with visual feedback of the transformation has a positive impact on the non-dominant hand's sEMG magnitude during the muscular-shaped avatar's embodiment. This partially supports **H2.1**. This positive was found superior to that of the passive transition with visual feedback. Therefore, this partially supports **H2.2**.

5.4.4 sEMG magnitude after embodiment

The Friedman test could not reveal significant differences in EVR_{after} for either hand across conditions (dominant: $p = .27$, non-dominant: $p = .83$). Therefore, I could not confirm hypothesis **H3.1** and **H3.2**.

Looking at the values of the EVR_{after}, it seems likely that either visual transition did not have any effect on the Proteus effects once the embodiment ended (see Fig. 5.8, bottom).

5.5 Discussion

In this chapter, I explored the use of visual effects to smooth the transition between the different stages of embodiment experiences in AR. This section summarizes and discusses our contributions to reaching this goal.

5.5.1 Transition system

I presented a proof-of-concept system applying visual effects when changing of embodiment perspective (1PP/3PP), when changing from one's real body to a generic avatar (3D mesh/regular-shaped avatar), when changing from this avatar's appearance to another one (regular/muscular), and the other way around. Using various techniques, these transitions were designed to allow for a more continuous embodiment experience than when abruptly switching perspectives or bodies, from one frame to another. Additionally, I explored two approaches to trigger transitions from one avatar to another: active transitions which are launched and synchronized with the user's action, and passive transitions which are applied automatically by the system and observed passively.

Although the transition system I present is new in its design, it was strongly inspired by previous work implementing embodiment systems in the 3PP. Like the system of Rosa et al. (2019), our setup allows seeing a remote avatar in the real world thanks to cameras embedded in the headset providing video see-through vision [66]. The main difference with their system is that, complementarily, I also use real-time 360 videos captured with another external camera to provide an "out-of-body" point-of-view where the real body is erased. This point of view is similar to what was done by Leggenhagger et al. (2007) [37]. Lastly, the 3D mesh capture of the user's real body was previously implemented by Nimchareon et al. (2018) who used a Microsoft Kinect to display the user's image in front of them in the real world, as I did [49]. Our system can be seen as an assemblage of these different setups, combining their features and allowing them to transition from one to another. Therefore, it contributes to demonstrating how existing techniques can be merged into an encompassing system, allowing users to flexibly adapt the embodiment experience.

I believe all of these transitions can benefit various kinds of embodiment setups, including VR ones. Contrary to VR where everything is virtual, AR embodiment requires integrating the avatar into the real environment with accuracy and coherence. As this cannot be done perfectly with current technologies yet, I feel that AR embodiment systems might have even further interest in having such transitions to mitigate the potential effects of discrepancies resulting from current system limits.

5.5.2 User study

To start evaluating the transition system I propose, I investigated the effects of one of its transitions on the SoE: the visual transitions applied between the embodiment of two avatars. I were interested in checking whether applying such transitions would allow for maintaining a stable SoE, even when the new avatar is more difficult to embody. To do so, I used our prototype system to assess such effects when switching from a regular-shaped avatar to a muscular one, inducing a lower SoE. I hypothesised that visual feedback showing the avatar's appearance progressively changing would help prevent the user's SoE from decreasing because the avatar's appearance gradually becomes different from the user's body. Further, I expected that visual feedback triggered by and synchronized with the user's physical action (F_ACTIVE) better help to retain the SoE than when visual feedback is presented automatically, without involving the user (F_PASSIVE).

Our results showed that applying active transitions with visual feedback synchronized with the user's actions helps to maintain the sense of agency. In contrast, applying an active transition without such visual feedback (N_ACTIVE) had the opposite effect as it reduced the sense of agency. I believe that users in the condition F_ACTIVE might not have felt in control of their avatar's transformation because they could not see how much their actions impacted it. This feeling of not controlling the transformation's progress may have in turn reduced the feeling of controlling the avatar itself. Therefore, when applying an active transition, visual feedback on the avatar's appearance change seems to be of primary importance to let the user feel in control of the transformation and ensure a positive effect of the transition. In line with this finding, some participants commented that they felt that they were controlling their avatars when they saw the avatar change synchronously with their squeezing actions.

Secondly, I were interested in verifying whether Proteus effect could benefit from visual transitions as a consequence of the higher stability of the SoE. I hypothesized that transitions with visual feedback would improve the intensity of the Proteus effect when and after embodying the muscular-shaped avatar, especially in the case of active transitions. Our results showed that an active transition combined with visual feedback partly led to stronger intensity of the Proteus effect during the embodiment (users applied more strength when squeezing the dumbbells). As the Proteus effect is linked to the strength of the SoE [27], this further supports the idea that condition F_ACTIVE was able to maintain a superior SoE than the other conditions. On the other hand, I could not observe a difference in the induced Proteus effect between our conditions after the end of the embodiment. Therefore, the impacts I observed on the Proteus Effect may have some limits: they might occur only during the embodiment and probably have short resilience over time. Further investigation is necessary to confirm this, but also to determine whether this is specific to our experiment and whether these effects can be reproduced with other kinds of avatars.

Enhanced impacts of Proteus effects influencing the user's physical performance can benefit various real-world tasks. For instance, in a hypothetical situation where a builder wears an AR headset while carrying heavy objects, transitioning from their real selves to a muscular avatar could provide them with stronger Proteus effects than when directly embodying this avatar. In addition, when people wish to do some exercise in the real world (e.g., in a park surrounded by magnificent nature), they can smoothly embody a muscular avatar standing in this physical location, thereby increasing their exercise performance through the intensified Proteus effect.

5.5.3 Limitations and future work

Although these results are interesting, they come with limitations that encourage follow-up research. First, I made the choice to use the 3PP to let users embody their avatars due to the difficulty of implementing the 1PP in AR. This gave us the opportunity to explore perspective transitions which I thought could highly benefit the continuity of the embodiment experience. However, previous research suggests that the 3PP tends to provide a weaker SoE than the 1PP [19, 22, 42–44, 57, 69, 72]. Therefore, there is a possibility that the effects I observed on the Proteus effects might have been more pronounced with this perspective.

To investigate the effects of visual transitions in the 1PP, future work will need to investigate how to achieve precise full-body tracking and alignment of the avatar on the user's body, but also ways to increase the vertical field of view of AR HMDs to allow seeing the avatar's body when looking down, as done with VR HMDs by Nakano et al. (2021) [48].

Secondly, although I present multiple types of transitions in our proof-of-concept system, our user study only evaluated active and passive transitions when changing from one avatar to another. The other transitions I provide (perspective transitions and switching from the real body mesh to the avatar) were experienced across all conditions and did not vary to avoid having too many variables to test. To investigate how they benefit the user experience and SoE, I encourage future work to evaluate their effects in dedicated studies.

Lastly, the present study evaluated the user's physical performance based on only one type of task (dumbbell squeezing) and relied on only one type of data (sEMG). Some of our unsupported hypotheses on the Proteus effects might have been supported by other types of physical tasks and measures such as heart rate or skin conductance values [20, 62]. Other limitations include differences by gender and age, which were not investigated. Although I recruited an equal number of male and female participants, I did not analyze the gender differences. In addition, participants were in their 20s and 30s, and I were not able to examine other age groups. In the future, follow-up studies would benefit from a wider demographic diversity.

5.6 Conclusions

In this chapter, I explored how visual transitions can benefit embodiment experiences in AR. I implemented a system allowing users to transition smoothly between embodiment perspectives, between their real body and an avatar, and between several avatars. I then conducted a user study to evaluate the effects of visual transitions when switching from a generic regular-shaped avatar to a muscular one on the SoE and Proteus effects. Two approaches were explored to trigger this transition: active transitions where the user is responsible for transforming their avatars, and passive transitions where the transformation is applied automatically. I also investigated the impact of displaying feedback showing the avatar transforming. To do so, I performed a cross-evaluation of two factors using the transition system I implemented: visual feedback (present/absent) and transition activation (active/passive). Our results partially supported that visual transitions with feedback help maintaining the SoE when switching to an avatar expected to induce a lower SoE (the muscular one). In addition, I discovered that active transitions with visual feedback retained the sense of agency better than passive transitions, but not when removing the visual feedback. Our results also partially supported the expectation that Proteus effects would also benefit from visual transitions during the embodiment of a muscular-shaped avatar, especially when the transition was active. This phenomenon could be exploited in future applications aiming to utilize the Proteus effects to facilitate real-world tasks. Therefore, I encourage future work to study whether these effects can be reproduced with other avatar appearances for different Proteus effects. Furthermore, it will be important that future research investigates the impact of changing an avatar's appearance using other types of physiological sensors than sEMG sensors, such as electroencephalogram (EEG) and electrodermal activity (EDA).

6 Discussion

In this section, I would describe discussion across the above three studies.

Through three studies, I investigated how AR avatar embodiment impacts the Proteus effect. Unlike VR, a user's real body is visible and it might be destructive to embody toward the avatar. Therefore, I dealt with a visible body by blurring or erasing it. In the first study, I addressed the effects of perspectives (1PP and 3PP) and avatar age (young and elder) on walking performance. The user's real body was blurred in the 3PP. In the second, the Proteus effect of a muscular-shaped avatar was studied. Further, in the third study I explored how visual transitions enhance the Proteus effect in AR. In the second and the third study, while a user embodied in an avatar, their own body was completely erased.

Three studies provided several findings and implications. For instance, I found in study #1 that a 3PP AR produces the similar amount of embodiment as a 1PP AR. This finding will encourage AR application developers to apply a 3PP for avatar embodiment. In addition, the results of study #2 supported the immediate and prolonged Proteus effect by a 3PP AR embodiment. Furthermore, study #3 demonstrated that visual transitions from a regular-built avatar to a muscular avatar enhanced the Proteus effect during embodiment. These findings might encourage AR application developers to apply a 3PP and visual transitions for avatar embodiment to improve user experience through enhanced Proteus effect.

However, there are some limitations in terms of the methodologies and user studies through studies. I would discuss the limitations and how to overcome them in future study.

6.1 First-person perspective

I could not investigate the Proteus effect using a 1PP in the ideal setting. I could not set up a 1PP embodiment where the user's own body was invisible and they could see the avatar moving in sync with their body in the real environment. Indeed, I compared a 1PP and a 3PP in study #1, but I used a static 360 image in stead of a real-time one. Moreover, in both study #2 and #3, a 3PP is applied to erase user own body perfectly and I did not study a 1PP embodiment. The result of study #1 showed that a 3PP in AR can induce the SoE as strong as a 1PP in AR. However, a 1PP avatar embodiment in the real environment should be addressed in future study since people use egocentric views to experience the physical world in their daily lives. In addition, many VR studies showed that a 1PP provides a stronger SoE compared with a 3PP [19, 42–44, 57, 69, 72]. Therefore, more research is needed to address the effects of a 1PP on the SoE and the Proteus effect in AR. To achieve the ideal 1PP AR embodiment, a real-time video inpainting technique would be required to precisely detect and erase the user's body.

6.2 Tracking limitation

In the presented systems, the avatar body movement was limited to avoid tracking errors. In study #1, the avatar did not replicate the user's finger motion. When users moved their fingers, in particular, closed their hands, they became aware of the difference in movements between the avatar and them by observing the motionless avatar's fingers. Therefore, as I mentioned in 3.5, some participants were not satisfied with an avatar's finger motion. In study #2 and study #3, in stead of tracking a whole body, a user's head and arms were tracked. The user's head was tracked by the HMD and their arms movement was replicated by the tracker mounted on the dumbbell. As a result, a user was required to keep on sitting while embodiment and their motion was limited to lifting up and down a dumbbell. Future research should explore the Proteus effect in AR where the user can move freely and the avatar moves in sync with their motion including any body parts such as fingers.

6.3 Experiment measures

There were limitations in the experiment measures. Through three studies, I focused on measuring the physical performance in the experiment. In study #1, the smoothness of walking and the walking speed were measured to examine the Proteus effect of an elder avatar. In study #2 and study #3, the physical task using dumbbells was conducted to figure out the impact of a muscular-shaped avatar. In study #2, the number of lift up-down the dumbbell and hand-grip strength were assessed. In study #3, sEMG sensors were employed to detect the user's muscular activity. By evaluating the physical performance, I could find the Proteus effect in AR. However, I could have measured how a user's perception would be changed by embodiment and visual transitions. For example, by asking participants how their perception of effort differed by embodying a muscular avatar, I could have found the potential effect of AR embodiment. If a muscular-shaped avatar has a positive impact on the Moreover, physiological sensors would allow to figure out the effect of embodiment. For example, EEG sensors might be used to detect how brain activity changes due to embodiment in future study.

6.4 Experiment participants

There were limitations in the experiment participants. First, the number of participants I could recruit was limited due to the sanitary restrictions of the Covid19 crisis in three studies. In study #1, I only collected eight participants. Moreover, in study #2, I only hired male participants due to recruiting limitation. Second, the age range of participants was biased toward the young generation. The average age in three user studies is 24.71 (study #1: $M = 25.75$, study #2: $M = 23.69$, study #3: $M = 24.94$). Future study needs to hire older participants to demonstrate that the Proteus effect is produced in AR in all generations. I expect that the positive impact of embodiment for older people would be practical in many fields. For example, the improved physical performance by a muscular-shaped avatar could be applied in the health-care field.

7 Conclusion

In this thesis, I explored how AR embodiment induces the Proteus effect and how the effect can be strengthened by visual transitions. The Proteus effect in AR could be used in many fields such as healthcare, education, and entertainment. For example, people might improve their physical performance in an exercise application by receiving the Proteus effect from a muscular avatar.

Nevertheless, AR embodiment and the Proteus effect has not been well studied, and so there is a need for more research in this area. This is mostly due to some technical challenges. First, in an AR applications the user's own body is visible in the real environment, and seeing both a real body and AR body might make the AR avatar embodiment less believable. Second, tracking errors can cause a misalignment between the user's real body and AR avatar which could deteriorate the AR embodiment experience. To address these challenges, I prepared a 3PP AR embodiment system where the user can view the avatar without any tracking errors. In the prototype AR system, the user's own body is blurred or completely erased. In addition, I used stable tracking devices to avoid tracking errors. Therefore, the user can focus on the avatar's appearance and observe that the avatar moves in sync with the user.

By using the prototype system, I conducted three studies to explore the effects of avatar appearance, perspectives, and avatar transitions on the Sense of Embodiement (SoE) and the Proteus effect. By doing this the thesis makes several research contributions; (1) it investigates the effect of perspective and avatar appearance in AR on user SoE and physical performance, (2) it examines the Proteus effect of a muscular avatar under fully erased user body conditions, and (3) it illustrates how the visual transitions influence the Proteus effect in AR.

Study #1 investigated the effects of perspectives (1PP versus 3PP) and avatar appearance (young versus elderly) on the SoE and the physical performance.

There were three hypotheses;

H1 There will be no significant difference between 1PP and 3PP conditions in the sense of embodiment as measured by the Peck SOE questionnaire [54].

H2 The smoothness of walking, as measured by the jerk cost, will be reduced in the elderly avatar condition, compared to the young avatar condition.

H3 The walking speed, measured by the walking time, becomes slower after the 3PP conditions, compared to 1PP conditions.

To evaluate the Proteus effect of an elderly looking avatar, I measured the smoothness of walking and the time to walk in a certain distance. The results showed that the 3PP embodiment induced the SoE as strong as the 1PP, and the 3PP decreased the smoothness of walking. Thus hypothesis H1 was confirmed, while H2 and H3 were rejected.

Study #2 examined the effects of avatar appearance (muscular versus non-muscular) on user SoE and the physical performance. In study #2, the user own body was completely erased, unlike study #1 where it was blurred. There were three hypotheses;

H1 Compared to the muscular avatar, participants would feel a higher sense of embodiment (SoE) for the non-muscular avatar.

H2 Compared to the non-muscular avatar, the muscular avatar would improve user's physical performance while controlling the avatar.

H3 Compared to the non-muscular avatar, the muscular avatar would improve user's physical performance after controlling the avatar.

I compared a muscular and a non-muscular avatar and measured the number of times of lifting the dumbbell and the hand-grip strength. The results supported the Proteus effect of a muscular avatar both during and after embodiment. Thus hypothesis H2 and H3 were confirmed, while H1 was rejected.

Study #3 investigated the effects of visual transitions (from similar-looking to different looking) on the SoE and the physical performance. I decomposed visual transitions into two factors (visual feedback (present/absent) and physical action (active/passive). In active transitions with visual feedback, an avatar's body gained muscular synchronized with the user's physical action, while the avatar's body changed automatically in passive transitions. In study #3 I employed sEMG sensors to measure user physical performance on a task that involved squeezing arms. There were six hypotheses;

H1-1 Compared to no feedback, visual feedback during transitions would maintain the SoE.

H1-2 Compared to passive transitions with visual feedback, active transitions with feedback would maintain the SoE.

H2-1 Compared to no feedback, visual feedback during transitions would enhance the Proteus effect during embodiment.

H2-2 Compared to passive transitions with visual feedback, active transitions with feedback would enhance the Proteus effect during embodiment.

H3-1 Compared to no feedback, visual feedback during transitions would enhance the Proteus effect after embodiment.

H3-2 Compared to passive transitions with visual feedback, active transitions with feedback would enhance the Proteus effect after embodiment.

Through the experiment, I confirmed that in the case of active transitions were held, visual feedback maintained the sense of agency better than no feedback. Additionally, active transitions with visual feedback enhanced the sense of agency and the Proteus effect during embodiment, compared to passive transitions with feedback. Thus hypothesis H1-1, H1-2, and H2-2 were partially confirmed, while H2-1, H3-1, and H3-2 were rejected.

Overall, results from the user studies with the prototype 3PP AR system indicates the potential of the Proteus effect in the real environment. They imply that future AR applications can utilize the diminishing technique and visual transitions in a 3PP which I used in the presented system. These techniques would let users feel more embodiment and receive the strong Proteus effect, then improve user experience.

However, there were some limitations with these studies. First, a 1PP embodiment in AR was not studied using the same setting as a 3PP where a user's real body is blurred or erased and an avatar moves in sync with the user. Secondly, the avatar body movement was limited to avoid tracking errors. Thirdly, there were limitations in the experiment measures. I could examine the Proteus effect in AR by evaluating a user's perception and physiological activity. Finally, there were limitations in the experiment participants due to recruiting limitation.

Accordingly, there are some important directions for future research that could be undertaken. More study is needed to confirm that a 1PP also induces the Proteus effect in the real environment. Although achieving avatar embodiment in a 1PP AR is difficult even with the latest technologies, future image processing and tracking methods could overcome the challenges of a 1PP AR embodiment. Secondly, a full body avatar tracking system could be applied to achieve an AR embodiment where the user can freely move their body. In addition, avatar transitions could be explored using avatars of different appearance. For instance, when the user embodies a tall avatar in the real environment, they might behave more confidently, based on a prior study in VR [91]. Furthermore, multiple physiological sensors might provide more insights into AR embodiment. Using sEMG sensors, the muscle activity was measured while the user was embodied in a muscular avatar. In a similar way EEG sensors could detect how user brain activity change through embodiment in an avatar in the real environment.

Overall, this thesis has opened the door to research of the Proteus effect in AR and I hope this work will inspire many more interesting studies in the future.

Acknowledgements

I would first like to thank my thesis supervisor Professor Kiyoshi Kiyokawa. If I had not met him, I would not jump into the VR/AR research world. He always helped me explore my master's degree. I cannot thank him enough. I would also like to thank Professor Hirokazu Kato as the second thesis supervisor. His opinions were always meaningful and critical. I would also like to thank Professor Hideaki Uchiyama. He gave me technical staff training. I would also like to thank Professor Naoya Isoyama. He taught me how to design experiments. I would also like to thank Professor Monica Perusquia-Hernandez. Her presentation on the experimental design was helpful. I would also like to thank Professor Anatole Lécuyer of the Hybrid research team, at Inria, the French National Institute for Research in Computer Science and Control, in Rennes, France. He constantly cheered me up to advance in my studies. I would also like to thank Professor Martin Hachet of the Potioc at Inria, in Bordeaux, France. He provided me with insightful comments. I would also like to thank PhD student Adélaïde Genay of the Potioc at Inria, in Bordeaux, France. I could not achieve the goal to present the paper without her help. I would also like to thank Mr. Yusuke Shikanai. He graduated from my university in 2021 and I took over his master's project. I would also like to acknowledge Professor Mark Billinghurst of the Empathic Computing Laboratory at the University of Auckland, New Zealand, and I am gratefully indebted to him for his very valuable feedback on this thesis. Through the internship program for one year, I acquired various beneficial knowledge. In particular, studying in Auckland was one of the most exciting experiences in my master's life. Finally, I must express my deepest gratitude to my parents and friends for their unwavering support and never-ending motivation during my years of study as well as during the process of conducting research for and writing this thesis. I appreciate from the bottom of my heart. Kia ora.

References

[1] D. Banakou, S. Kishore, and M. Slater. Virtually being einstein results in an improvement in cognitive task performance and a decrease in age bias. *Frontiers in psychology*, p. 917, 2018.

[2] M. E. Benalcázar, C. Motoche, J. A. Zea, A. G. Jaramillo, C. E. Anchundia, P. Zambrano, M. Segura, F. B. Palacios, and M. Pérez. Real-time hand gesture recognition using the myo armband and muscle activity detection. *2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM)*, pp. 1–6, 2017.

[3] M. Botvinick and J. Cohen. Rubber hands 'feel'touch that eyes see. *Nature*, 391(6669):756–756, 1998.

[4] R. G. Buckley, C. R. Stehman, F. L. Dos Santos, R. H. Riffenburgh, A. Swenson, N. Mjos, M. Brewer, and S. Mulligan. Bedside method to estimate actual body weight in the emergency department. *The Journal of Emergency Medicine*, 42(1):100–104, 2012. doi: 10.1016/j.jemermed.2010.10.022

[5] D. Burin, K. Kilteni, M. Rabuffetti, M. Slater, and L. Pia. Body ownership increases the interference between observed and executed movements. *PloS one*, 14(1):e0209899, 2019.

[6] Y.-W. Cha, H. Shaik, Q. Zhang, F. Feng, A. State, A. Ilie, and H. Fuchs. Mobile. egocentric human body motion reconstruction using only eyeglasses-mounted cameras and a few body-worn inertial sensors. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 616–625, 2021. doi: 10.1109/VR50410.2021.00087

[7] S. Cmentowski, A. Krekhov, and J. Krüger. Outstanding: A multi-perspective travel approach for virtual reality games. In *Proceedings of*

the annual symposium on computer-human interaction in play, pp. 287–299, 2019.

[8] H. G. Debarba, E. Molla, B. Herbelin, and R. Boulic. Characterizing embodied interaction in first and third person perspective viewpoints. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 67–72, 2015. doi: 10.1109/3DUI.2015.7131728

[9] A. Denisova and P. Cairns. First person vs. third person perspective in digital games: Do player preferences affect immersion? In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, p. 145–148. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2702123.2702256

[10] D. Dewez, R. Fribourg, F. Argelaguet, L. Hoyet, D. Mestre, M. Slater, and A. Lécuyer. Influence of personality traits and body awareness on the sense of embodiment in virtual reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 123–134, 2019. doi: 10.1109/ISMAR.2019.00-12

[11] E. Ebrahimi, L. S. Hartman, A. Robb, C. C. Pagano, and S. V. Babu. Investigating the effects of anthropomorphic fidelity of self-avatars on near field depth perception in immersive virtual environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1–8, 2018. doi: 10.1109/VR.2018.8446539

[12] K. Emmerich, A. Krekhov, S. Cmentowski, and J. Krueger. Streaming vr games to the broad audience: A comparison of the first-person and third-person perspectives. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445515

[13] T. Feuchtner and J. Müller. Extending the body for interaction with reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, p. 5145–5157. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3025453.3025689

[14] R. Fribourg, F. Argelaguet, A. Lécuyer, and L. Hoyet. Avatar and sense of embodiment: Studying the relative preference between appearance, control and point of view. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):2062–2072, 2020. doi: 10.1109/TVCG.2020.2973077

[15] H. Galvan Debarba, S. Bovet, R. Salomon, O. Blanke, B. Herbelin, and R. Boulic. Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality. *PLOS ONE*, 12(12):1–19, 12 2017. doi: 10.1371/journal.pone.0190109

[16] A. C. S. Genay, A. Lecuyer, and M. Hachet. Being an avatar "for real": a survey on virtual embodiment in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 2021. doi: 10.1109/TVCG.2021.3099290

[17] G. Gorisse, O. Christmann, E. A. Amato, and S. Richir. First- and third-person perspectives in immersive virtual environments: Presence and performance analysis of embodied users. *Frontiers in Robotics and AI*, 4, 2017. doi: 10.3389/frobt.2017.00033

[18] G. Gorisse, O. Christmann, S. Houzangbe, and S. Richir. From robot to virtual doppelganger: Impact of visual fidelity of avatars controlled in third-person perspective on embodiment and behavior in immersive virtual environments. *Frontiers in Robotics and AI*, 6, 2019. doi: 10.3389/frobt.2019.00008

[19] A. Guterstam and H. H. Ehrsson. Disowning one's seen real body during an out-of-body illusion. *Consciousness and Cognition*, 21(2):1037–1042, 2012. Standing on the Verge: Lessons and Limits from the Empirical Study of Consciousness. doi: 10.1016/j.concog.2012.01.018

[20] A. Halbig and M. E. Latoschik. A systematic review of physiological measurements, factors, methods, and applications in virtual reality. *Frontiers in Virtual Reality*, 2, 2021. doi: 10.3389/frvir.2021.694567

[21] N. Hogan. An organizing principle for a class of voluntary movements. *J Neurosci*, 4(11):2745–2754, 1984. doi: 10.1523/JNEUROSCI.04-11-02745

[22] M. Hoppe, A. Baumann, P. C. Tamunjoh, T.-K. Machulla, P. W. Woźniak, A. Schmidt, and R. Welsch. There is no first- or third-person view in virtual reality: Understanding the perspective continuum. In *CHI Conference on Human Factors in Computing Systems (CHI '22)*, 2022. doi: 10.1145/3491102.3517447

[23] M. Jung, J. Kim, and K. Kim. Measuring recognition of body changes over time: A human-computer interaction tool using dynamic morphing and body ownership illusion. *PLOS ONE*, 15(9):1–14, 2020. doi: 10.1371/journal.pone.0239322

[24] S. Jung, G. Bruder, P. J. Wisniewski, C. Sandor, and C. E. Hughes. Over my hand: Using a personalized hand in vr to improve object size estimation, body ownership, and presence. SUI '18, p. 60–68. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3267782.3267920

[25] S. Jung, P. J. Wisniewski, and C. E. Hughes. In limbo: The effect of gradual visual transition between real and virtual on virtual body ownership illusion and presence. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 267–272, 2018. doi: 10.1109/VR.2018.8447562

[26] M. Kari, T. Grosse-Puppendahl, L. F. Coelho, A. R. Fender, D. Bethge, R. Schütte, and C. Holz. Transformr: Pose-aware object substitution for composing alternate mixed realities. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 69–79, 2021. doi: 10.1109/ISMAR52148.2021.00021

[27] K. Kilteni, I. Bergstrom, and M. Slater. Drumming in immersive virtual reality: The body shapes the way we play. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):597–605, 2013. doi: 10.1109/TVCG.2013.29

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

[29] K. Kilteni, A. Maselli, K. P. Kording, and M. Slater. Over my fake body: body ownership illusions for studying the multisensory basis of own-body

perception. *Frontiers in Human Neuroscience*, 9, 2015. doi: 10.3389/fnhum.2015.00141

[30] D. Kim, S. Woo, J.-Y. Lee, and I. S. Kweon. Deep video inpainting. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5792–5801, 2019.

[31] D. Kim, S. Woo, J.-Y. Lee, and I. S. Kweon. Recurrent temporal aggregation framework for deep video inpainting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(5):1038–1052, 2020.

[32] M. Kocur, F. Habler, V. Schwind, P. W. Woźniak, C. Wolff, and N. Henze. Physiological and perceptual responses to athletic avatars while cycling in virtual reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445160

[33] M. Kocur, M. Kloss, V. Schwind, C. Wolff, and N. Henze. *Flexing Muscles in Virtual Reality: Effects of Avatars' Muscular Appearance on Physical Performance*, p. 193–205. Association for Computing Machinery, New York, NY, USA, 2020.

[34] M. Kocur, D. Roth, and V. Schwind. Towards an investigation of embodiment time in virtual reality. In C. Hansen, A. Nürnberger, and B. Preim, eds., *Mensch und Computer 2020 - Workshopband*. Gesellschaft für Informatik e.V., Bonn, 2020. doi: 10.18420/muc2020-ws134-339

[35] S. Lee, S. W. Oh, D. Won, and S. J. Kim. Copy-and-paste networks for deep video inpainting. In *International Conference on Computer Vision (ICCV)*, 2019.

[36] B. Lenggenhager, M. Mouthon, and O. Blanke. Spatial aspects of bodily self-consciousness. *Consciousness and Cognition*, 18(1):110–117, 2009. doi: 10.1016/j.concog.2008.11.003

[37] B. Lenggenhager, T. Tadi, T. Metzinger, and O. Blanke. Video ergo sum: Manipulating bodily self-consciousness. *Science*, 317(5841):1096–1099, 2007. doi: 10.1126/science.1143439

[38] Z. Li, C.-Z. Lu, J. Qin, C.-L. Guo, and M.-M. Cheng. Towards an end-to-end framework for flow-guided video inpainting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[39] J.-H. T. Lin, D.-Y. Wu, and J.-W. Yang. Exercising with a six pack in virtual reality: Examining the proteus effect of avatar body shape and sex on self-efficacy for core-muscle exercise, self-concept of body shape, and actual physical activity. *Frontiers in Psychology*, 12, 2021. doi: 10.3389/fpsyg.2021.693543

[40] M. R. Longo, F. Schüür, M. P. Kammers, M. Tsakiris, and P. Haggard. What is embodiment? a psychometric approach. *Cognition*, 107(3):978–998, 2008. doi: 10.1016/j.cognition.2007.12.004

[41] J.-L. Lugrin, M. Landeck, and M. Latoschik. Avatar embodiment realism and virtual fitness training. 03 2015. doi: 10.1109/VR.2015.7223377

[42] A. Maselli and M. Slater. The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience*, 7, 2013. doi: 10.3389/fnhum.2013.00083

[43] A. Maselli and M. Slater. Sliding perspectives: dissociating ownership from self-location during full body illusions in virtual reality. *Frontiers in Human Neuroscience*, 8, 2014. doi: 10.3389/fnhum.2014.00693

[44] D. Medeiros, R. K. dos Anjos, D. Mendes, J. a. M. Pereira, A. Raposo, and J. Jorge. Keep my head on my shoulders! why third-person is bad for navigation in vr. VRST '18. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3281505.3281511

[45] G. Morais, L. Neves, A. Masiero, and M. C. Castro. Application of myo armband system to control a robot interface. pp. 227–231, 01 2016. doi: 10.5220/0005706302270231

[46] S. Mori, S. Ikeda, and H. Saito. A survey of diminished reality: Techniques for visually concealing, eliminating, and seeing through real objects. *IPSJ Transactions on Computer Vision and Applications*, 9:1–14, 2017.

[47] A.-A. Nabulsi. *HAND GESTURE RECOGNITION VIA ELECTROMYOGRAPHIC (EMG) ARMBAND FOR CAD SOFTWARE CONTROL*. PhD thesis, 01 2018. doi: 10.13140/RG.2.2.14996.12167

[48] K. Nakano, N. Isoyama, D. Monteiro, N. Sakata, K. Kiyokawa, and T. Narumi. Head-mounted display with increased downward field of view improves presence and sense of self-location. *IEEE Transactions on Visualization & Computer Graphics*, 27(11):4204–4214, 2021.

[49] C. Nimcharoen, S. Zollmann, J. Collins, and H. Regenbrecht. Is that me?—embodiment and body perception with an augmented reality mirror. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 158–163, 2018. doi: 10.1109/ISMAR-Adjunct.2018.00057

[50] J.-M. Normand, E. Giannopoulos, B. Spanlang, and M. Slater. Multisensory stimulation can induce an illusion of larger belly size in immersive virtual reality. *PLOS ONE*, 6(1):1–11, 01 2011. doi: 10.1371/journal.pone.0016128

[51] N. Ogawa, T. Narumi, and M. Hirose. Virtual hand realism affects object size perception in body-based scaling. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 519–528, 2019. doi: 10.1109/VR.2019.8798040

[52] S. Osimo, R. Pizarro, B. Spanlang, and M. Slater. Conversations between self and self as sigmund freud—a virtual body ownership paradigm for self counselling open. *Scientific Reports*, 5, 10 2015. doi: 10.1038/srep13899

[53] A. Oyanagi and R. Ohmura. Transformation to a bird: Overcoming the height of fear by inducing the proteus effect of the bird avatar. In *Proceedings of the 2nd International Conference on Image and Graphics Processing*, ICIGP '19, p. 145–149. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3313950.3313976

[54] T. C. Peck and M. Gonzalez-Franco. Avatar embodiment. a standardized questionnaire. *Frontiers in Virtual Reality*, 1, 2021. doi: 10.3389/frvir.2020.575943

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

[56] M. Perusquía-Hernández, F. Dollack, C. K. Tan, S. Namba, S. Ayabe-Kanamura, and K. Suzuki. Smile action unit detection from distal wearable electromyography and computer vision. In *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*, pp. 1–8, 2021. doi: 10.1109/FG52635.2021.9667047

[57] V. Petkova, M. Khoshnevis, and H. H. Ehrsson. The perspective matters! multisensory integration in ego-centric reference frames determines full-body ownership. *Frontiers in Psychology*, 2, 2011. doi: 10.3389/fpsyg.2011.00035

[58] V. I. Petkova and H. H. Ehrsson. If i were you: Perceptual illusion of body swapping. *PLOS ONE*, 3(12):1–9, 12 2008. doi: 10.1371/journal.pone.0003832

[59] J. S. Pierce, R. Pausch, C. B. Sturgill, and K. D. Christiansen. Designing a successful hmd-based experience. *Presence: Teleoperators and Virtual Environments*, 8(4):469–473, 1999. doi: 10.1162/105474699566350

[60] A. Pomes and M. Slater. Drift and ownership toward a distant virtual body. *Frontiers in Human Neuroscience*, 7, 2013. doi: 10.3389/fnhum.2013.00908

[61] M. Pyasik, G. Tieri, and L. Pia. Visual appearance of the virtual hand affects embodiment in the virtual hand illusion. *Scientific Reports*, 10, 03 2020. doi: 10.1038/s41598-020-62394-0

[62] A. Raheel, M. Majid, M. Alnowami, and S. M. Anwar. Physiological sensors based emotion recognition while experiencing tactile enhanced multimedia. *Sensors*, 20(14), 2020. doi: 10.3390/s20144037

[63] R. Ratan, D. Beyea, B. J. Li, and L. Graciano. Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: A meta-analysis of the proteus effect. *Media Psychology*, 23(5):651–675, 2020. doi: 10.1080/15213269.2019.1623698

[64] R. Reinhard, K. G. Shah, C. A. Faust-Christmann, and T. Lachmann. Acting your avatar's age: effects of virtual reality avatar embodiment on real life walking speed. *Media Psychology*, 23(2):293–315, 2020. doi: 10.1080/15213269.2019.1598435

[65] H. Rhodin, C. Richardt, D. Casas, E. Insafutdinov, M. Shafiei, H.-P. Seidel, B. Schiele, and C. Theobalt. Egocap: Egocentric marker-less motion capture with two fisheye cameras. *ACM Trans. Graph.*, 35(6), dec 2016. doi: 10.1145/2980179.2980235

[66] N. Rosa, J.-P. van Bommel, W. Hürst, T. Nijboer, R. C. Veltkamp, and P. Werkhoven. Embodying an extra virtual body in augmented reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1138–1139, 2019.

[67] R. S. Rosenberg, S. L. Baughman, and J. N. Bailenson. Virtual superheroes: Using superpowers in virtual reality to encourage prosocial behavior. *PLOS ONE*, 8(1):1–9, 01 2013. doi: 10.1371/journal.pone.0055003

[68] A. Salagean, J. Hadnett-Hunter, D. J. Finnegan, A. A. De Sousa, and M. J. Proulx. A virtual reality application of the rubber hand illusion induced by ultrasonic mid-air haptic stimulation. *ACM Trans. Appl. Percept.*, 19(1), jan 2022. doi: 10.1145/3487563

[69] S. Seinfeld and J. Müller. Impact of visuomotor feedback on the embodiment of virtual hands detached from the body. *Scientific Reports*, 10, 02 2020.

[70] Y. Shikanai, N. Isoyama, N. Sakata, and K. Kiyokawa. Phantact: An augmented reality system to induce the proteus effect in the real world. *MVE IPSJ-CVIM*, 120(319):1–6, 2021.

[71] M. Slater, D. Pérez Marcos, H. Ehrsson, and M. Sanchez-Vives. Towards a digital body: the virtual arm illusion. *Frontiers in Human Neuroscience*, 2, 2008. doi: 10.3389/neuro.09.006.2008

[72] M. Slater, B. Spanlang, M. V. Sanchez-Vives, and O. Blanke. First person experience of body transfer in virtual reality. *PLOS ONE*, 5(5):1–9, 05 2010. doi: 10.1371/journal.pone.0010564

[73] M. Slater, A. Steed, J. McCarthy, and F. Marinelli. The virtual ante-room: Assessing presence through expectation and surprise. 1998.

[74] A. Smolentsev, J. E. Cornick, and J. Blascovich. Using a preamble to increase presence in digital virtual environments. *Virtual Reality*, 21:153–164, feb 2017. doi: 10.1007/s10055-017-0305-4

[75] D. Sproll, J. Freiberg, T. Grechkin, and B. E. Riecke. Poster: Paving the way into virtual reality - a transition in five stages. In *2013 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 175–176, 2013. doi: 10.1109/3DUI.2013.6550235

[76] F. Steinicke, G. Bruder, K. Hinrichs, M. Lappe, B. Ries, and V. Interrante. Transitional environments enhance distance perception in immersive virtual reality systems. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, pp. 19–26, 2009.

[77] F. Steinicke, G. Bruder, K. Hinrichs, and A. Steed. Special section: Ieee virtual reality (vr) 2009: Gradual transitions and their effects on presence and distance estimation. *Comput. Graph.*, 34(1):26–33, feb 2010. doi: 10.1016/j.cag.2009.12.003

[78] F. Steinicke, G. Bruder, K. Hinrichs, A. Steed, and A. L. Gerlach. Does a gradual transition to the virtual world increase presence? In *2009 IEEE Virtual Reality Conference*, pp. 203–210. IEEE, 2009.

[79] K. Suzuki, F. Nakamura, J. Otsuka, K. Masai, Y. Itoh, Y. Sugiura, and M. Sugimoto. Recognition and mapping of facial expressions to avatar by embedded photo reflective sensors in head mounted display. In *2017 IEEE Virtual Reality (VR)*, pp. 177–185, 2017. doi: 10.1109/VR.2017.7892245

[80] K. Tatarian, M. Couceiro, E. Ribeiro, and D. Faria. Stepping-stones to transhumanism: An emg-controlled low-cost prosthetic hand for academia. 09 2018. doi: 10.1109/IS.2018.8710489

[81] G. Tieri, A. Gioia, M. Scandola, E. F. Pavone, and S. M. Aglioti. Visual appearance of a virtual upper limb modulates the temperature of the real hand:

a thermal imaging study in immersive virtual reality. *European Journal of Neuroscience*, 45(9):1141–1151, 2017.

[82] G. Tieri, G. Morone, S. Paolucci, and M. Iosa. Virtual reality in cognitive and motor rehabilitation: facts, fiction and fallacies. *Expert review of medical devices*, 15(2):107–117, 2018.

[83] G. Tieri, E. Tidoni, E. F. Pavone, and S. M. Aglioti. Mere observation of body discontinuity affects perceived ownership and vicarious agency over a virtual hand. *Experimental brain research*, 233(4):1247–1259, 2015.

[84] M. Tsakiris and P. Haggard. The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of experimental psychology. Human perception and performance*, 31 1:80–91, 2005.

[85] M. Tsakiris, G. Prabhu, and P. Haggard. Having a body versus moving your body: How agency structures body-ownership. *Consciousness and Cognition*, 15(2):423–432, 2006. doi: 10.1016/j.concog.2005.09.004

[86] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1643–1652, 2018. doi: 10.1109/TVCG.2018.2794629

[87] M. L. Weijs, E. Macartney, M. M. Daum, and B. Lenggenhager. Development of the bodily self: Effects of visuomotor synchrony and visual appearance on virtual embodiment in children and adults. *Journal of Experimental Child Psychology*, 210:105200, 2021. doi: 10.1016/j.jecp.2021.105200

[88] E. Wolf, N. Döllinger, D. Mal, C. Wienrich, M. Botsch, and M. E. Latoschik. The embodiment of photorealistic avatars influences female body weight perception in virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 462–473, 2020. doi: 10.1109/ISMAR50242.2020.00071

[89] E. Wolf, M. L. Fiedler, N. Döllinger, C. Wienrich, and M. E. Latoschik. Exploring presence, avatar embodiment, and body perception with a holographic augmented reality mirror. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 350–359, 2022. doi: 10.1109/VR51125.2022.00054

[90] W. Xu, A. Chatterjee, M. Zollhöfer, H. Rhodin, P. Fua, H.-P. Seidel, and C. Theobalt. Mo2cap2: Real-time mobile 3d motion capture with a cap-mounted fisheye camera. *IEEE transactions on visualization and computer graphics*, PP, 03 2018. doi: 10.1109/TVCG.2019.2898650

[91] N. Yee and J. Bailenson. The proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, 33:271–290, 2007. doi: 10.1111/j.1468-2958.2007.00299.x

[92] Y. Yuan and A. Steed. Is the rubber hand illusion induced by immersive virtual reality? In *2010 IEEE Virtual Reality Conference (VR)*, pp. 95–102, 2010. doi: 10.1109/VR.2010.5444807

8 Publication

[1] R. Otono, A. C. S. Genay, M. Perusquía-Hernández, N. Isoyama, H. Uchiyama, M. Hachet, A. Lécuyer and K. Kiyokawa, "I'm Transforming! Effects of Visual Transitions to Change of Avatar on the Sense of Embodiment in AR," the 2023 IEEE Conference on Virtual Reality and 3D User Interfaces (VR 2023), 2023. (accepted)

[2] R. Otono, A. C. S. Genay, M. Perusquía-Hernández, N. Isoyama, H. Uchiyama, M. Hachet, A. Lécuyer and K. Kiyokawa, "Studying "Avatar Transitions" in Augmented Reality: Influence on Sense of Embodiment and Physiological Activity," 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), 2022, pp. 503-504, doi: 10.1109/ISMAR-Adjunct57072.2022.00106. (published)

[3] R. Otono, N. Isoyama, H. Uchiyama and K. Kiyokawa, "Third-Person Perspective Avatar Embodiment in Augmented Reality: Examining the Proteus Effect on Physical Performance," 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 2022, pp. 730-731, doi: 10.1109/VRW55335.2022.00216. (published)

[4] R. Otono, Y. Shikanai, K. Nakano, N. Isoyama, H. Uchiyama and K. Kiyokawa, "The Proteus Effect in Augmented Reality: Impact of Avatar Age and User Perspective on Walking Behaviors," The 26th Annual Conference of the Virtual Reality Society of Japan (VRSJ), 2021, 1C3-1 (published)

[5] T. Fujisawa, R. Otono. T. Sasaki, K. Miyazaki, N. Isoyama, H. Uchiyama and K. Kiyokawa, "Fingun," The 26th Annual Conference of the Virtual Reality Society of Japan (VRSJ), 2021, IVRC-16 (published)

9 Appendix

Table 9.1: The items of SoE self-report based on [54].

Item number	Question
Q1	At some point it felt that the virtual body resembled my own (real) body, in terms of shape, skin tone or other visual features.
Q2	I felt as if the movements of the virtual body were influencing my own movements.
Q3	At some point it felt as if my real body was starting to take on the posture or shape of the virtual body that I saw.
Q4	It seemed as if my feet was touching the floor seen in HMD. (in Study #1) / It seemed as if my hands was touching the dumbbell. (in Study #2)
Q5	I felt out of my body.
Q6	I felt as if my body was located where I saw the virtual body.
Q7	It seemed as if I felt the touch of the floor in the location where I saw the virtual feet touched. (in Study #1) / It seemed as if I felt the touch of the dumbbell in the location where I saw the virtual hands touched. (in Study #2)
Q8	It felt as if my (real) body were turning into an 'avatar' body.
Q9	I felt like I could control the virtual body as if it was my own body.
Q10	I felt like I was wearing different clothes from when I came to the laboratory.
Q11	I felt as if my body had changed.
Q12	I felt a realistic sensation in my body when I saw the virtual hand.
Q13	It seemed as if the touch I felt was caused by the floor touching the virtual feet. (in Study #1) / It seemed as if the touch I felt was caused by the dumbbell touching the virtual hands. (in Study #2)
Q14	I felt as if the virtual body was my body.
Q15	I felt as if my (real) body were drifting toward the virtual body or as if the virtual body were drifting toward my (real) body.
Q16	I felt that my own body could be affected by the world seen in HMD.