Being a Virtual Avatar in a Real Environment : Design and Study of Avatar Embodiment in Augmented Reality

Adélaïde Genay

université de BORDEAUX

université de BORDEAUX

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR

DE L'UNIVERSITÉ DE BORDEAUX

ECOLE DOCTORALE MATHÉMATIQUES ET INFORMATIQUE

SPÉCIALITÉ INFORMATIQUE

Par **Adélaïde GENAY**

Être un Avatar Virtuel dans un Environnement Réel : Conception et Étude de l'Incarnation d'Avatars en Réalité Augmentée

Sous la direction de : **Martin HACHET**
Co-directrice : **Anatole LÉCUYER**

Soutenue le 13 Décembre 2022

Membres du jury :

Mme Catherine PELACHAUD, Directrice de recherche, Institut des Systèmes Intelligents et de Robotique (ISIR) (Paris), Rapporteure
M. Frank STEINICKE, Professeur, Université de Hambourg (Hambourg, ALLEMAGNE), Rapporteur
Mme Mavi SANCHEZ-VIVES, Professeur, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS) (Barcelona, ESPAGNE), Examinatrice
M. Jean-Louis VERCHER, Directeur de recherche, CNRS PARIS (Paris), Examinateur
M. Pascal GUITTON, Professeur émérite, Université de Bordeaux (Talence), Président
M. Martin HACHET, Directeur de recherche, Inria Bordeaux Sud-Ouest, LaBRI, CNRS, Université de Bordeaux (Talence), Directeur de thèse
M. Anatole LECUYER, Directeur de recherche, Inria Rennes Bretagne-Atlantique, IRISA (Rennes), Co-Directeur

Titre : Être un Avatar Virtuel dans un Environnement Réel : Conception et Étude de l'Incarnation d'Avatars en Réalité Augmentée

Résumé : Autrefois réservés aux environnements virtuels, les avatars sont désormais de plus en plus présents en Réalité Augmentée (RA). Cette technologie permet d'interagir avec du contenu virtuel directement intégré à notre environnement physique. Associée aux techniques de suivi corporel, elle permet de se voir à l'intérieur d'un corps virtuel de n'importe quelle apparence et d'étendre nos capacités d'action. Malgré les nombreuses possibilités d'applications, peu de recherches ont exploré comment la mettre en œuvre et l'exploiter. De plus, la façon dont les utilisateurs perçoivent de telles expériences dans des environnements réels n'est pas encore bien comprise. Dans cette thèse, nous explorons l'incarnation des avatars en RA au travers de plusieurs axes. Tout d'abord, nous introduisons une taxonomie pour mieux décrire les expériences d'incarnation de par un continuum d'avatarisation corporelle. Ensuite, à l'aide d'un affichage optique de RA, nous étudions le sentiment d'incarnation de l'utilisateur vis-à-vis de mains virtuelles et l'impact du mélange du virtuel et du réel sur ce sentiment. En troisième lieu, nous examinons les cas d'utilisation de l'incarnation d'avatars en RA et étudions les techniques d'interaction permettant de les contrôler à la troisième personne pour explorer le monde réel. Nous proposons une mise en œuvre de ce concept et évaluons l'approche d'utiliser un tel système pour améliorer notre perception des affordances du monde réel au travers d'une étude utilisateurs. Pour terminer, nous explorons des stratégies pour améliorer les illusions d'incarnation à l'aide de transitions graphiques et des méthodes pour préparer l'utilisateur à son "avatarisation".

Mots-Clés: Avatar, Réalité Augmentée, Réalité Mixte, Interactions Human-Machine

---

Title: Being a Virtual Avatar in a Real Environment: Design and Study of Avatar Embodiment in Augmented Reality

Abstract: Once reserved for virtual environments, avatars are now increasingly present in Augmented Reality (AR). This technology allows interacting with virtual content directly integrated into our physical surroundings. Combined with body tracking, it makes it possible to experience being inside a virtual body of any desired appearance and to extend one's action capabilities. Despite the numerous opportunities for applications that this new possibility provides, little research has explored how to implement and exploit it. The way users perceive such experiences within real environments is additionally not well understood yet. In this thesis, we explore the embodiment of avatars in AR through several axes. First, we introduce a taxonomy through a body avatarization continuum to better describe embodiment experiences. Next, using an optical see-through AR display, we study the user's sense of embodiment towards virtual hands and how mixing the virtual with the real impacts this sense. We then look at the use cases for avatar embodiment in AR and investigate interaction techniques allowing us to control avatars in the third person and to explore the real world. We propose an implementation of this concept and evaluate the approach of using such a system to improve our perception of real-world affordances through a user study. Finally, we explore strategies to enhance embodiment illusions by designing transitions to become one's avatar and methods to prepare the user for their "avatarization".

Keywords: Avatar, Augmented Reality, Mixed Reality, Human-Computer Interaction

---

Equipe POTIOC Inria/LABRI

UMR 5800 Université de Bordeaux, 33000 Bordeaux, France.

Contents

List of Figures

List of Tables

List of Publications

Journals

- Adélaïde Genay, Anatole Lécuyer, and Martin Hachet (2021a). "Being an Avatar "for Real": a Survey on Virtual Embodiment in Augmented Reality". In: *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1. DOI: 10.1109/TVCG.2021.3099290
- Adélaïde Genay, Anatole Lécuyer, and Martin Hachet (2021b). "Virtual, Real or Mixed: How Surrounding Objects Influence the Sense of Embodiment in Optical See-Through Experiences?" In: *Frontiers in Virtual Reality* 2. DOI: 10.3389/frvir.2021.679902

Conference papers

- Adélaïde Genay, Anatole Lécuyer, and Martin Hachet (Oct. 2022). "What Can I Do There? Controlling AR Self-Avatars to Better Perceive Affordances of the Real World". In: *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. Singapour. DOI: 10.1109/ISMAR55827.2022.00061
- Riku Otono, Adélaïde Genay, Monica Perusquía-Hernández, Naoya Isoyama, Hideaki Uchiyama, Martin Hachet, Anatole Lécuyer, and Kiyoshi Kiyokawa (Mar. 2023). "Active or Passive Transition? Effects of Avatar Transitions on the Sense of Embodiment and Physiological Activity". In: *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Shanghai, China. URL: https://hal.inria.fr/hal-03968274/

Posters and workshops

- Adélaïde Genay, Anatole Lécuyer, and Martin Hachet (June 2020). "Incarner un Avatar en Réalité Augmentée : Revue de la Littérature". In: *Workshop sur les Affects, Compagnons artificiels et Interactions (WACAI)*. CNRS, Université Toulouse Jean Jaurès, Université de Bordeaux. Saint Pierre d'Oléron, France. URL: https://hal.inria.fr/hal-02933477
- Riku Otono, Adélaïde Genay, Monica Perusquía-Hernández, Naoya Isoyama, Hideaki Uchiyama, Martin Hachet, Anatole Lécuyer, and Kiyoshi Kiyokawa (Oct. 2022). "Studying "Avatar Transitions" in Augmented Reality: Influence on Sense of Embodiment and Physiological Activity". In: *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. Singapour. DOI: 10.1109/ISMAR-Adjunct57072.2022.00106

Résumé long en Français

Context

Le sentiment d'incarnation, tel que décrit par Kilteni, Groten, et al. (2012), est composé de trois sous-composantes : l'agentivité (le sentiment d'être l'auteur des mouvements d'un corps), la propriété corporelle (le sentiment que le corps est la source des sensations ressenties) et l'auto-localisation (le sentiment d'être situé à l'intérieur du corps). Ces sentiments sont toujours présents et semblent empêcher la dissociation de notre corps à nous-mêmes. Pourtant, il a été démontré que l'application d'une stimulation sensorielle appropriée pouvait supprimer cette barrière et conduire à l'illusion d'avoir un autre corps. L'illusion de la main en caoutchouc de Botvinick et al. (1998) est l'illustration la plus célèbre de ce phénomène, produisant la sensation de posséder une main en caoutchouc par une stimulation visuotactile synchrone. De nombreuses autres expériences ont suivi celle-ci en l'étendant à des mannequins entiers puis à des avatars virtuels (Kilteni, Maselli, et al., 2015).

Les possibilités élargies offertes par les dispositifs de Réalité Virtuelle (RV) ont conduit cette technologie à être largement impliquée dans la recherche sur le sentiment d'incarnation et ont permis d'identifier de nombreux facteurs l'influençant. Parmi eux, il a été montré que l'apparence de l'avatar, les interactions avec des éléments virtuels et les retours sensoriels affectent de manière importante les illusions de propriété corporelle. Des facteurs ascendants tels que les traits de personnalité de l'utilisateur ont également été associés à la capacité de faire l'expérience d'illusions d'incarnation.

Les progrès en Réalité Augmentée (RA) ont amené les chercheurs à utiliser de plus en plus cette technologie pour explorer ces sujets également. Lorsqu'ils sont incarnés en RA, les avatars permettent aux utilisateurs de se voir dans le monde réel mais dans un corps de différente forme, taille ou apparence. Cette possibilité peut être pertinente dans de nombreux cas, que ce soit dans l'éducation, les jeux-vidéos, la collaboration à distance, le soin médical ou simplement comme moyen d'expression (Y. Wang et al., 2022). Cependant, encore peu de choses sont connues sur la façon dont le mélange d'éléments virtuels et réels interagit avec ces illusions et on ne sait pas si elles diffèrent de celles vécues dans des mondes entièrement réels ou entièrement virtuels. Étant donné la prévalence croissante

des technologies de RA et le nombre grandissant d'applications offrant des superpositions virtuelles à leurs utilisateurs, la recherche sur les mécanismes d'incarnation dans de tels environnements visuels semble plus importante que jamais.

Démarche

L'objectif de cette thèse de doctorat a été d'étendre les recherches précédentes sur l'incarnation des avatars en les amenant en dehors des environnements virtuels grâce aux technologies de RA. Nos travaux ont suivi les trois axes de recherche suivants :

RQ1 **Facteurs d'influence du sentiment d'incarnation en RA**

Un aspect qui rend la RA fondamentalement différente de la RV est la présence de contenus virtuels et réels mélangés. Ainsi, la question suivante se pose naturellement : le contraste notable qui existe entre le monde réel et le contenu virtuel a-t-il un impact sur la façon dont les utilisateurs perçoivent leurs avatars en RA ? Nous tentons de répondre à cette question dans la première partie de cette thèse.

RQ2 **Utilisations de l'incarnation d'avatars en RA**

Le fait que la RA préserve la vision du monde réel ouvre la porte à de nouveaux domaines d'application pour les expériences d'incarnation. Elle rend possible de changer d'apparence corporelle, mais aussi de voir une copie virtuelle de soi-même d'un point de vue extérieur. Ce "détachement" virtuel permet d'explorer des environnements distants sans avoir à y entrer physiquement. En deuxième partie de cette thèse, nous proposons d'étudier comment concevoir un système pour une telle utilisation et comment celle-ci peut bénéficier à notre perception de l'espace physique.

RQ3 **Lisser l'expérience utilisateur de l'avatarisation en RA**

Dans la plupart des systèmes d'incarnation, les utilisateurs découvrent le corps de leurs avatars au moment même où ils mettent leur casque. Ce changement brutal peut quelque peu dérouter, surtout lorsqu'on passe directement à l'incarnation d'un corps virtuel à la troisième personne comme proposé dans l'axe de recherche précédent. Cette observation soulève donc des questions supplémentaires : comment le passage de l'état de soi à celui d'avatar peut-il être vécu de manière plus naturelle ? Nous proposons d'étudier des solutions à cette question dans la dernière partie de cette thèse de doctorat.

Pour restreindre notre recherche, nous avons concentré notre intérêt sur les affichages optiques transparents de RA et sur les avatars en tant que représentations virtuelles 3D anthropomorphiques des utilisateurs.

Contributions

Notre objectif principal a été de mieux comprendre comment les utilisateurs perçoivent les avatars en RA, comment cette perception peut leur être utile dans des scénarios de la vie quotidienne, et comment renforcer les illusions d'incarnation. Pour atteindre cet objectif, les contributions suivantes ont été réalisées :

- Nous avons proposé une nouvelle taxonomie pour les expériences d'incarnation des avatars en s'appuyant sur un continuum "d'avatarisation" du corps.
- Nous avons étudié l'impact du mélange de contenus virtuels et réels sur les illusions d'incarnation en RA.
- Nous avons proposé et mis en œuvre le concept de contrôler des avatars à distance pour explorer le monde réel qui nous entoure.
- Nous avons évalué comment la manipulation de tels avatars peut améliorer notre perception des affordances physiques.
- Nous avons conçu et testé des stratégies pour faciliter les expériences d'avatarisation.

Avant de discuter des perspectives apportées par cette thèse, la section suivante résume brièvement comment nous avons réalisé ces différentes contributions.

Résumé des contributions

Notre recherche a commencé par un travail théorique visant à mieux comprendre et décrire les expériences d'incarnation : en Chapitre 1, nous avons présenté un **continuum d'avatarisation corporelle** offrant un moyen de catégoriser les systèmes d'incarnation en fonction de l'étendue de la représentation virtuelle de l'utilisateur. À l'aide de ce continuum, nous avons ensuite passé en revue la littérature sur la RA mettant en œuvre l'incarnation d'avatars en Chapitre 2. Cela a permis de clarifier les aspects techniques et les défis liés à l'avatarisation du corps en RA, les méthodes disponibles pour induire et mesurer le sentiment d'incarnation, ainsi que les similitudes existantes entre l'incarnation en RA et en RV. D'autre part, nous avons identifié un manque de recherche sur la façon dont le contenu affiché en RA influence la perception de l'utilisateur de ses avatars.

En Chapitre 3, nous avons approfondi ce sujet par le biais d'une **étude utilisateur visant à déterminer si le contenu mixte des environnements de RA pouvait affecter les illusions d'incarnation**. Notre expérience a comparé le sentiment d'incarnation de l'utilisateur envers des mains de robot virtuelles incarnées en présence de différentes quantités d'objets virtuels et réels. Nous avons développé un système d'affichage transparent permettant aux participants de voir ces mains superposées à leurs propres mains. Nos résultats montrent que les participants ont ressenti une

plus forte appropriation des mains virtuelles lorsqu'elles étaient visionnées en présence d'objets virtuels et réels mélangés, par rapport au cas où les mains virtuelles étaient le seul contenu virtuel visible. Cela suggère que le contenu de l'environnement devrait être pris en compte lors de la conception d'expériences d'incarnation. Cependant, nous n'avons pas pu conclure avec certitude sur l'origine des différences observées et leur étendue à des populations plus importantes. Cette expérience devrait néanmoins encourager la communauté à approfondir l'idée que l'avatar lui-même n'est pas le seul modérateur du sentiment d'incarnation, mais que le contexte et le support de l'incarnation peuvent également jouer un rôle.

Ensuite, nous avons étudié comment l'avatarisation corporelle et la RA peuvent être combinées pour des cas d'utilisation concrets. En puisant dans les théories sur la perception des affordances, nous avons proposé **le concept d'envoyer son avatar pour agir à distance et améliorer notre compréhension du monde physique**, ainsi que de nos possibilités d'action au sein de celui-ci. Le Chapitre 4 détaille la preuve de concept que nous avons mise en place pour illustrer cette idée. Le système développé permet de manipuler un avatar dont la forme et la taille correspondent à celles de l'utilisateur. L'avatar peut être contrôlé à l'aide de trois techniques : avec une manette, avec des interactions manuelles et avec un suivi corporel. Nous avons décrit des scénarios où chacune de ces techniques peut être pertinente et discuté de leurs forces et faiblesses individuelles.

Le Chapitre 5 a poursuivi ces recherches en **évaluant les effets du contrôle d'avatars à la troisième personne sur la perception des affordances du monde réel**. Nous avons mené trois études durant lesquelles des participants ont contrôlé des avatars (personnalisés à leur taille) pour tester des actions dans le monde réel. La première expérience a montré que la manipulation d'une version virtuelle de soi peut effectivement améliorer la précision et la confiance de l'utilisateur lorsqu'il ou elle estime ses capacités action. La deuxième expérience a étudié comment la mémoire de la manipulation de l'avatar peut être exploitée lors de la réitération des mêmes jugements une fois que l'avatar n'est plus en vue. Nous n'avons pas obtenu de preuve que le fait d'avoir utilisé un avatar avait un impact sur les performances de l'utilisateur, car nous n'avons pas trouvé de différence significative entre les conditions qui incluaient une utilisation antérieure de l'avatar et celles qui ne l'incluaient pas. La dernière expérience a poursuivi notre enquête sur les effets résilients des avatars en demandant aux utilisateurs d'estimer leur aptitude à passer une porte coulissante. Nous avons également comparé leur précision face à une porte réelle et face à une porte virtuelle. Nous n'avons pas trouvé d'effets significatifs de l'avatar, ni de la virtualité de la porte sur la performance de l'utilisateur. Néanmoins, de la première étude, nous retenons que l'expérience acquise à travers l'avatar peut être utilisée pour prendre des décisions plus éclairées et pour mieux saisir la dimension de notre environnement.

Enfin, nous avons étudié des **stratégies pour renforcer les illusions d'incarnation et réduire la variabilité** causée par les caractéristiques individuelles des utilisateurs. Le Chapitre 6 décrit notre recherche collaborative avec le laboratoire CARE (à Nara, au Japon) sur la façon d'exploiter les différentes phases des expériences d'incarnation, et en particulier, le temps passé *avant* l'incarnation. Nous avons présenté une étude d'utilisateur évaluant comment la préparation mentale de l'utilisateur peut aider à créer des illusions plus fortes. L'expérience a été menée en RV avec un avatar représentant Hulk, un personnage de la franchise Marvel. Avant de l'incarner, les participants ont réalisé un exercice d'imagination et d'actorat que nous avons conçu afin de les aider à se mettre dans la peau de leur avatar. Les résultats de cette expérience étaient encore en cours d'analyse lors de l'écriture de ce manuscrit. Les premiers tests ne montrent pas d'effet significatif de cette préparation mentale sur le sentiment d'incarnation. Cependant, nous avons pu constater des effets positifs sur l'expérience utilisateur globale: les participants se sont sentis plus fort et étaient moins conscients de l'environnement autour d'eux. D'autres analyses seront menées après la soumission de cette thèse.

Conclusion

Il reste beaucoup à faire avant d'obtenir des avatars de la trempe de ceux des œuvres de science-fiction. Les technologies actuelles de RA permettent de se voir à l'intérieur d'un corps virtuel et de le contrôler dans un espace physique, mais elles ne permettent que des interactions limitées et ne fournissent qu'un retour tactile rudimentaire. Par conséquent, la prochaine grande étape semble être la création d'environnements multisensoriels qui impliquent non seulement l'intégration réaliste d'avatars dans le monde réel, mais aussi des sensations de toucher lorsqu'on les incarne. Nous encourageons la poursuite des recherches dans cette direction, car elles peuvent contribuer à améliorer à la fois les illusions d'incarnation et l'expérience générale des utilisateurs.

Deuxièmement, des progrès importants doivent être réalisés dans la recherche sur le contrôle des avatars. Des avancées récentes permettent aujourd'hui de reproduire les mouvements des utilisateurs sur leurs avatars en temps réel, sans avoir à porter d'équipement. Cependant, l'agentivité reste limitée par le nombre de points qu'il est possible de suivre simultanément et la précision du suivi reste insuffisante pour fournir une animation de l'avatar fluide. De plus, les casques de RA actuels étant difficilement compatibles avec le suivi facial, les avatars ne peuvent que rarement représenter les expressions de leur utilisateur correctement.

Malgré le nombre important de recherches qu'il reste à faire, la prochaine génération d'avatars n'est plus très loin. Ainsi, il semble nécessaire de continuer les recherches permettant d'en comprendre la perception et d'anticiper les enjeux.

One of the things our grandchildren will find
quaintest about us is that we distinguish the digital
from the real, the virtual from the real.
In the future, that will become literally impossible.

William Gibson

1

Introduction

This chapter describes the motivation to complete this thesis. Inspired by philosophy, our goal is to explore the perceptive and interactive aspects of virtual avatars embodied in Augmented Reality.

1.1 The Avatar Dream

The term "soma-sêma" refers to one of the most well-known mysteries taught by Pythagoras and Plato in ancient Greece. It is a simple wordplay that translates to "body-tomb" and exposes the antiquated idea that the body is the prison of the mind (Irigaray, 2017). Indeed, it is clear that the body inherently limits what we can do and who we can be. One cannot suddenly decide to fly and just fly, or become a cat and turn into one. Current society further extends these physical limits into social limits as it usually imposes an identity defined by body characteristics. Attributing an identity in such ways becomes particularly problematic when one's physical identity does not match who one feels they are, such as in the case of trans people.

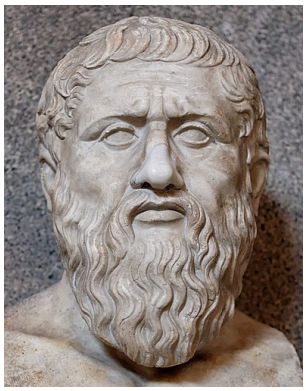


Figure 1.1: Herm representing Plato. The first known expression of the wordplay "soma-sêma" can be found in *Gorgias* (492e–493a, transl. W. R. M. Lamb).

The fantasy of being freed from the social and physiological limitations of the body has been the heart of many works of science fiction. Within these, avatars are often envisioned as a solution to overcome them. Neal Stephenson's 1992 novel "Snow Crash" describes virtual avatars as technologies to reimagine the self: "Your avatar can look any way you want it to, up to the limitations of your equipment. If you're ugly, you can make your avatar beautiful." The best-seller of Ernest Cline, "Ready Player One", is another example of fiction where avatars have become an interface to escape reality. And of course, the blockbuster "Avatar" by James Cameron features a human able to surpass his physical disability (and even death) by transferring his mind into the body of a Na'vi, an ultimate form of an avatar.

The vision that virtual avatars can offer the means to transcend the physical limits of the human body is not reserved for fiction. It is also shared by academics such as Harrell et al. (2017), who see it as being one of society's major technological dreams that they define as follows:

> **The Avatar Dream**
> A culturally shared vision of a future in which, through the computer, people can become whoever or whatever they want to be.
>
> Harrell et al. (2017)

As explained in their article, the Avatar Dream comes alongside other technological dreams. For example, the idea of ubiquitous computing is the fantasy of everything becoming smart, including phones, houses, or automobiles. The dream of artificial intelligence is one of sentient, conscious machines. In contrast to these, the Avatar Dream aims to transcend the corporeality that is tied to our biological condition through virtuality.

The term *avatar* comes from the Sanskrit word *avatarah*. It originally stands for the "descent" of a Hindu god into our world, under a terrestrial form. Today, this term is used in various ways but it usually designates a virtual character representing a user within a virtual environment (Bailenson et al., 2004). They are often distinguished from *agents*, which are the digital representation of artificial intelligence. In this manuscript, we will use the general definition of the Berkshire encyclopedia of human-computer interactions:

Definition Avatar

"Perceptible digital representations whose behaviors reflect those executed, typically in real-time, by a specific human being."

Bailenson et al. (2004)

Avatars can faithfully represent the user or not (e.g. user scan, generic human model, or even other creatures), but also be visually realistic or stylized. Some researchers see this as an opportunity to virtually acquire "superpowers", as virtual content can take any shape and is not subject to physical laws (Sadeghian et al., 2021).

Beyond extending our physical abilities, Harrell et al. (2017) argue that achieving the Avatar Dream would allow people to gain empowerment when exploring and constructing their self-image. Indeed, avatars make it possible to see oneself within another gender's body, with different skin color, or physical condition. Virtual embodiment could therefore help people to negotiate the social constraints that limit who they can be in the real world but also increase their empathy towards people that are different from them.

Although featured in recent fiction works, technologies that enable the Avatar Dream have already started to permeate our society. Online Virtual Reality (VR) communities are successful examples of their implementation. These communities let users explore alternative versions of themselves within immersive virtual environments

and live completely different experiences through the first-person embodiment of customizable avatars[1].

However, virtual communities are limited to virtual environments at this point. Upon removing their headsets, users return to their real bodies and lose the privileges of their virtual bodies. The avatar dream is therefore tied to another of society's technological dreams, that of the "Ultimate Display".

[1] VRChat is a popular example of such social platforms. Like in Stephenson's novel, one may experience what it is to embody a different gender, have new abilities such as flying, or even be another creature.

1.2 *Augmented Reality, the "Ultimate Display"*

In his essay "The Ultimate Display", I. Sutherland (1965) speculates on the future of computers, screens, and human-computer interaction. His article was visionary for new technology, predicting the arrival of computer keyboards, mice, touchscreens, the bucket fill tool (e.g. on Paint), but also of Extended Reality (XR) displays. Describing the latter, he wrote:

> The Ultimate Display would, of course, be a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal. With appropriate programming, such a display could literally be the Wonderland into which Alice walked.
>
> Ivan E. Sutherland (1965)

The pursuit of the Ultimate Display is driven by the desire for a means to perfectly trick the human senses into believing something illusory is real. Hence, achieving it may be the key to achieving the Avatar Dream, as it would provide a way to deceive the user into perceiving a virtual body as their own. Currently, it seems that our dreams may be the closest experience we have of such deception. The imaginary worlds and sensations we perceive through them feel undoubtedly real, even when being irrational. However, most of us are unable to control our dreams and they remain individual experiences; in this way, they cannot measure up to the vision of the Ultimate Display (Bown et al., 2017).

Ivan Sutherland saw the potential for technology to provide the experience of "dreaming while awake". Three years after publishing his essay, he presented the first prototype for the Ultimate Display which he called the "Sword of Damocles" (see Figure 1.2). The author's initial description of the Ultimate Display seemed somewhat similar to that of current CAVE systems (Cruz-Neira et al., 1992). Yet, upon presenting his prototype, it seems that what he had in mind was to augment our surrounding reality rather than create a new one. Indeed, the system was not a room so to speak, but a head-mounted display through which the user could observe 3D models in the real world thanks to cathode-ray tubes (one for each eye) and optical elements. Users had to be strapped to the device to create a better immersion, which made the experience rather uncomfortable. This device was the first Augmented Reality (AR) display.

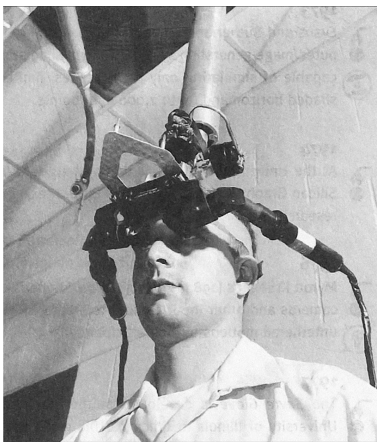


Figure 1.2: The "Sword of Damocles" (I. E. Sutherland, 1968). This AR device was released before even the invention of personal computers. Its nickname came from the fact that it was extremely heavy and needed to be suspended from the ceiling, above the user's head. Still regarded as a significant step in the development of AR, this work along with other pioneering contributions in computer graphics led Ivan Sutherland to receive the Turing Award in 1988.

Naturally, the term *augmented reality* did not exist at this point. It arrived much later in the 1990s when two scientists (P.C. Thomas and W.M. David) developing an experimental AR system for Boeing finally coined it. Multiple researchers have worked on definitions of this concept since then. For this manuscript, we will adopt one of the most widely accepted definitions of AR:

Definition **Augmented Reality (AR)**

"Systems that have the following three characteristics:

1. Combines the real and virtual,
2. Interactive in real-time,
3. Registered in 3-D"

Azuma (1997)

To situate AR among other technologies, we additionally rely on the seminal essay of Milgram and Kishino (1995). They position AR and VR along the same continuous scale which they called the *reality-virtuality continuum*. In their view, the physical world would reside on one of the extremes of this scale, and fully virtual spaces on the opposite extreme. Regions in between describe other realms combining different degrees of reality and virtuality, including AR, but also Augmented Virtuality (AV) where physical content appears inside of immersive virtual space. Altogether, this continuum defines the larger encompassing concept of Mixed Reality (MR).

Rather than immersing users in a new virtual world, AR enhances the one they live in. This makes its experience very different from that of VR and provides it with a clear advantage: it allows taking advantage of the flexibility of virtuality right into one's physical surroundings. AR is therefore seen as a promising solution for the shortcomings of current embodiment systems (e.g. VR social platforms), which confine the experience of avatars to virtual environments. Through AR, users could continuously benefit from their avatars at home, in public spaces, or even at work.

On the other hand, implementing virtual embodiment in AR is also more challenging. Indeed, AR systems require building an awareness of the physical environment's geometry for the virtual content to be properly integrated. Implementing realistic embodiment experiences is consequently more difficult in AR where the user's real body remains visible by default. This visibility implies that the avatar needs to be aligned with its user in real time with sufficient responsiveness, robustness, and precision for the illusion of being inside another body to be plausible.

Despite important progress, AR devices of today are not quite yet the ones envisioned by Ivan Sutherland; they do not produce strong essential copies of things, and the content displayed is not realistic and responsive enough to be indistinguishable from physical objects. Yet, their experience approaches that of the Ultimate Display. Equipment and applications initially confined to a few professional fields (design offices, research labs, etc.) are starting to emerge through-

out society. Off-the-shelf AR devices of today are not only capable of building spatial maps of physical space in real-time, but also of tracking user gestures to enable free-hand direct interaction with 3D holograms integrated into the real world. The addition of supplementary components pushes even farther such capabilities: although still in a premature form, AR already allows the embodiment of full-bodied avatars animated by one's body movements when coupled with state-of-the-art tracking devices.

This unprecedented possibility of embodying another body within a physical environment opens the door for countless new experiences, with applications beyond what VR currently offers in video games, cinema, therapy, education, communication, and self-expression. Motivated by these new opportunities, the rapid growth of AR research and industry encourages the idea that this technology will soon reach the same level of usability as VR (Masood et al., 2019). The general interest in AR has become particularly striking since the announcements of tech giants on their intent to develop *metaverses*. This concept refers to immersive, self-sustaining, and shared MR spaces. As put into words by Y. Wang et al. (2022), they are "created by the convergence of physically persistent virtual space and virtually enhanced physical reality". The development of avatars and AR technologies is key to achieving this merge.

However, whereas research activities in computer vision have considerably increased to meet the needs for the robustness of AR devices, how users perceive AR content remains an under-explored question, and even more so when it comes to avatars. Investigating the impacts of AR avatars on their users, therefore, seems to have become more important than ever.

1.3 *Virtual self-modeling: the effects of avatars*

Research conducted over the past 20 years has revealed how flexible our notion of "my body" is: applying the right sensory input at the right time is sufficient to create the illusion of having another body, known as a Body Ownership Illusion (BOI). The Rubber Hand Illusion (RHI) of Botvinick et al. (1998) was one of the first and most famous illustrations of this phenomenon, producing the feeling of owning a rubber hand through visuotactile stimulation.

Many variations of their experiment have appeared since then, extending it to full-body mannequins (Valeria I. Petkova et al., 2008) and eventually to virtual avatars (Kilteni, Maselli, et al., 2015). They rely on the user's brain to interpret the sensory stimuli presented as originating from the observed fake body, and not from the real body. When this succeeds, the user may experience the feeling of owning this body, called the Sense of Embodiment (SoE).

This sense was described in the research of Kilteni, Groten, et al. (2012) as being composed of three sub-components: the sense of *self-location* which is the feeling of being located inside the body, the sense of *ownership* corresponding to the feeling that the body is the

source of experienced sensations, and the sense of *agency* which is the feeling of being the author of a body's movements (see Figure 1.3). In this work, we will use the following working definition:

Definition Sense of Embodiment (SoE)

"The SoE toward a body B is the sense that emerges when B's properties are processed as if they were the properties of one's own biological body."

Kilteni, Groten, et al. (2012)

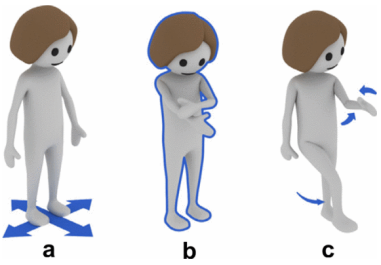


Figure 1.3: The sense of self-location (a), body ownership (b), and agency (c). Some also consider the sense of change to be an important part of the SoE. This sense refers to the change perceived in the body schema due to the stimulation. Figure extracted from the work of Roth and Latoschik (2020).

Since users cannot see their bodies in VR, giving them an avatar provided a natural solution to integrate them into virtual worlds. For this reason, VR displays were involved early on in the research on the SoE of avatars. Immersive environments allowed testing the influence of various factors on the SoE that could not be investigated before. Among these factors, the avatar's appearance (Schwind et al., 2017), interaction with virtual elements (Argelaguet et al., 2016) or sensory feedback (Fribourg et al., 2020) were shown to affect ownership illusions importantly. Bottom-up factors such as user personality traits were also found to have a play in the user's ability to experience embodiment (e.g. locus of control, sensory suggestibility) (Dewez et al., 2019; Marotta et al., 2016).

As the understanding of how avatars are perceived evolved, the central question of whether embodying a virtual body has effects on its users was raised. Could the characteristics of particular avatars lead us to act and express ourselves differently, or even to perceive our world differently? Driven by these interrogations, Yee et al. (2007) demonstrated that people in virtual worlds sometimes unconsciously adapt their behavior to match their avatars. For instance, participants to whom they had assigned taller avatars behaved more confidently and aggressively during negotiation tasks, whereas those with more attractive avatars behaved more intimately with other individuals. They came to the conclusion that the participants altered their conduct to match the behavior they thought other people would expect their digital self-representation to have. Proteus, a Greek god who had the power to assume the appearance of any entity, was the inspiration behind Yee and Bailenson's decision to name this phenomenon the *Proteus Effect*.

Research has since then found that specific avatar characteristics could provoke numerous beneficial outcomes, including the reduction of implicit racial bias (Peck, Seinfeld, et al., 2013) and body dissatisfaction (Fox et al., 2013), or the improvement of math performance (Banakou, Kishore, et al., 2018) and creative thinking (Guegan et al., 2016). The meta-analysis of Ratan et al. (2020) provides a comprehensive overview of the effects that were observed and their respective strengths. On the other hand, the Proteus effect can also generate negative behaviors such as self-objectification or hostility (Fox et al., 2013; Hawkins et al., 2021). The current watchword for developers is therefore to remain cautious when designing immersive

social platforms (Praetorius et al., 2020). Nevertheless, the opportunities for positive self-transformation that avatars provide are seen as worth pushing by many researchers, foreseeing their practical fields of application. Recent articles notably already encourage exploiting avatars for behavior change strategies during various medical therapies (eating disorders, pain treatment, etc.)[2].

[2] A complete review can be found in the work of Rheu et al. (2020) who surveyed the literature proposing health interventions using avatars.

The benefits of transposing such application scenarios to AR are easy to conceive. The possibility that avatars could generate different effects inherent to real-world contexts also raises a great deal of interest. For example, one could imagine using avatars to enhance the perception of real environments.

However, research investigating the effects and applications of avatar embodiment was almost exclusively conducted in immersive VR and little is known about them in AR contexts. AR remains fundamentally different from VR, and generalizing previous results to AR without verification seems unwise. Therefore, before investigating the existence of phenomena such as the Proteus effect, the first step is to understand the relationship between users and their avatars in environments where the real and the virtual are mixed. One of the goals of this thesis was to contribute to this objective.

1.4 *Proposing a Body Avatarization Continuum*

AR enables virtual body augmentations and transfigurations at different levels, ranging from the simple addition of virtual body accessories to a complete change of appearance and morphology. These virtual augmentations further enlarge the already complex phenomenology of embodiment experiences. In his attempt to categorize them, De Vignemont (2011) stressed the need for a taxonomy capable of capturing this diversity.

We introduce a body avatarization continuum to address the manifold of virtual embodiment and to clarify what we consider "being an avatar" is. Inspired by the reality-virtuality continuum, we describe the scale of the virtual user representation as a linear expansion, presented in Figure 1.4.

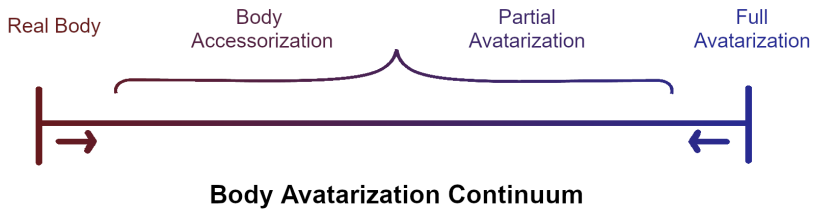


Figure 1.4: The body avatarization continuum covers the extent of the user's virtual representation within a virtual embodiment system, regardless of the environmental context.

At the extreme left of the continuum, the user deals only with their **real body**, deprived of any virtual element. This representation corresponds to the natural physical appearance of users in their daily lives. At the opposite end is what we pinned as **full avatarization**. The user embodies and controls a complete virtual body. This

body can be visualized either in the first person through the eyes of an avatar, or in the third person as in out-of-body experiences. In the latter case, considering such experience as full avatarization may seem paradoxical, especially when the user's real body remains visible. However, previous research suggests that an SoE can occur towards disconnected virtual avatars as well (J. v. Bommel, 2017; Lenggenhager et al., 2007; Rosa, J.-P. v. Bommel, et al., 2019). We, therefore, have included them in this area of the continuum.

The region between the two extrema relates to diverse degrees of virtual body augmentation. **Body accessorization** describes the addition of superficial elements such as virtual clothes or glasses to the user's body. Examples of such augmentations can be found in the projector-camera system of Harrison et al. (2011) where a watch is displayed on the user's wrist. **Partial avatarization**, on the other hand, involves the addition of virtual body parts or removal of real ones as well as the modification of bodily characteristics. This includes limbs, skin color, face shape, and other features of the human anatomy. For instance, the supernumerary hand illusion implemented by Gilbers (2017) corresponds to partial avatarization as it provides the user with an additional limb — a third hand — integrated into their real body. [3]

The question of where body augmentation starts being avatarization is another complex subject, falling within the realm of epistemology. On the one hand, removing or placing accessories on our bodies will not change the philosophical notion of "me" (Parfit, 1971). On the other hand, when choosing attire, we assume a particular social identity that elicits corresponding behaviors linked to our image of society (Adam et al., 2012; Stone, 1962). Indeed, it seems that clothes and accessories do not only impact the way others perceive and judge us, but also the way we see ourselves. A significant number of studies seem to indicate that this self-image has quite important consequences on our behavior, often in unconscious ways (K. Johnson et al., 2014). For instance, people wearing formal clothes might want to convey an image of professionalism and formality, but also feel more focused and confident in negotiating with others (Kraus et al., 2014; Slepian et al., 2015).

This top-down influence named "enclothed cognition" (Adam et al., 2012) is reminiscent of the embodiment of avatars: our intrinsic identity remains the same, but wearing clothes and accessories broadcasts a self-representation that we control and identify to. As an effort to resolve ambiguity, this article assumes that virtual body accessorization eventually amounts to embodying a hybrid avatar, i.e. a representation mixing real and virtual elements.

In the following sections, the term "body avatarization" will be used indistinctly to refer to the implementation of any of the virtual embodiment experiences defined by the body avatarization continuum.

[3] Whether the modification of traits such as hair length, eye color, etc. corresponds to partial avatarization or body accessorization is up for discussion. Being customarily modified in real life (e.g. through hair extensions, contact lenses), we decided to place such changes between these two areas.

1.5 Contributions of this Ph.D.

The questions at the core of this Ph.D. thesis were part of a collaborative research project called AVATAR, conducted from 2018 to 2022 by the French national research institute Inria. The AVATAR project aimed at delivering the next generation of virtual selves in digital worlds. It sought to push further the limits of perception and interaction by developing new avatars that are better embodied, more interactive, and more social. This ambitious goal implied tackling three great scientific challenges: (i) improving how we acquire, model and simulate avatars with novel capacities, (ii) designing and evaluating novel 3D interaction paradigms for avatar-based interaction, and (iii) designing and evaluating new multi-sensory feedback to better feel the subsequent interactions through avatars.

In order to address these multidisciplinary challenges, several research teams with complementary expertise were involved. I was part of the HYBRID and POTIOC teams, involved in the project for their strong know-how in human computer-interaction, multisensory feedback, and neuroscience[4]. My input in this project was to explore the embodiment of avatars in AR, the subject of my thesis.

[4] Other Inria teams that were involved were MORPHEO for their expertise in reconstruction, MIMETIC for avatar animation and biomechanics, GRAPHDECO for rendering, and MJOLNIR for input devices

1.5.1 Directions, scope, and rationale

The aim of this Ph.D. thesis was to extend previous research on the virtual embodiment of avatars by taking it outside of virtual environments through AR technologies. As mentioned in the previous sections, much remains to be addressed before we reach an understanding of avatar embodiment in AR comparable to the one attained in VR. Therefore, our research focused on continuing to scratch the surface of this subject with the following three research axes:

RQ1 **Factors of influence of the SoE in AR**

One aspect that makes AR fundamentally different from VR is the presence of virtual and real content mixed together. In VR, the user's vision is completely overridden by computer-generated graphics, and the avatar is perfectly merged into the virtual environment. In AR, on the other hand, the avatar usually stands out from the physical environment as being of a different nature. Telling apart virtual content from real one remains easy even with state-of-the-art displays due to rendering limitations. A question that naturally arises is therefore the following: does this noticeable contrast between the real world and virtual content impact the way users perceive their avatars in AR? We attempt to answer this question in the first part of this thesis.

RQ2 **Usages of avatar embodiment in AR**

The fact that AR preserves the real world's vision also opens the door to new areas of application for embodiment experiences. In particular, AR can provide the flexibility of virtual bodies while allowing interaction with people and the surrounding physical

environment. Hence, it is possible to change the body appearance but also enables one to witness a virtual copy of themselves from a third-person point of view. One could imagine using this ability to explore distant environments without physically entering them. We propose to investigate how to design a system for such an application and how it may benefit our perception of physical space.

RQ3 **Smoothing the user experience of body avatarization in AR**

In most embodiment systems, users see the body of their avatars as soon as they put their headsets on. This abrupt change may be somewhat confusing when directly switching from being in one's body to embodying a virtual body in the third person, as proposed in the previous research axis. This observation, therefore, raises additional questions: how can the change from being oneself to being an avatar be experienced more smoothly? We propose to investigate solutions for this issue as the last part of this Ph.D. thesis.

To narrow our research, we focused our interest on optical see-through AR displays[5] and on avatars as anthropomorphic 3D virtual representations of users.

[5] These displays integrate computer-generated visuals in the direct line of sight of a user's eyes through transparent optical combiners.

Our motivation to study avatars in AR can be summarized with the following prospects, inspiring our research: in the future, (i) embodying avatars in AR could allow extending one's **action** possibilities by replacing our body which capacities are limited with a digital one. (ii) They may also empower users when exploring and expressing their **identity** by enabling costless, unlimited, and reversible changes in their appearance. Combined with AR, they make it possible to deal with social constraints of who one can be in public contexts. Finally, (iii) avatar embodiment may become a powerful tool for extending **cognition**. Through AR, it could be used in real-life situations beyond what VR currently offers, e.g. to enhance one's cognitive abilities for precise needs or correct biased perceptions.

Of course, like all technologies, avatars come with their share of ethical questioning. Without being blind to the utopian nature of the perspectives listed above, the work presented in this thesis looks optimistically in their direction. The goals of our research are therefore necessarily anticipatory: we seek to create a better understanding of how avatar embodiment is experienced for the design of future systems to take into account. We additionally explore applications for AR avatars and investigate how to maximize their interests, while keeping in sight the ideals described above.

1.5.2 Contributions and outline of this manuscript

Our contributions to the research questions listed before are organized into seven chapters, including this one. Figure 1.5 offers an overview of this manuscript's outline.

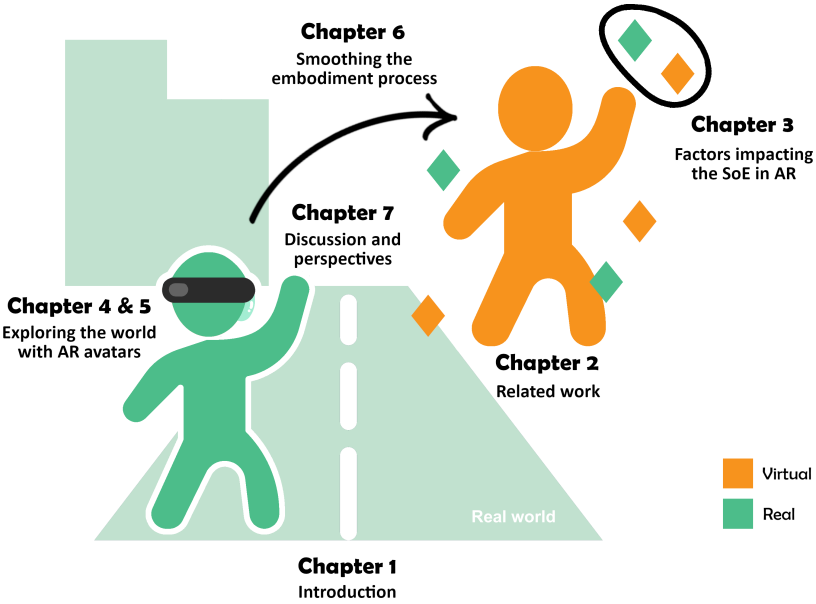


Figure 1.5: Illustration of the outline of this Ph.D. thesis. Additional information is placed in the appendices as follows: **Appendix A.1**: evaluation material used in the experiments of Chapter 5 and their results. **Appendix B**: instructions and questionnaires applied in the experiment of Chapter 6.

Chapter 2. We begin with a literature review on the embodiment of virtual avatars in AR. Our survey summarizes how body avatarization was previously used in AR applications at each area of the continuum. Then, it provides an overview of the technical challenges linked to AR embodiment systems. We review the choices available when implementing such a system, including the user perspective, display type, animation technique, and 3D registration method. We also describe the strategies previously used to induce and measure the SoE in AR. Lastly, we discuss how this previous literature is connected to our research axes.

Chapter 3. As a second step, we investigated the effects of AR environments on BOIs. Our goal was to evaluate whether mixing real and virtual content had an impact on the SoE (RQ1). To do so, we built an OST fishtank system and ran a study where users controlled virtual robot hands superimposed on their real hands within three AR environments, containing progressive amounts of virtual objects. Our results suggest that the BOI was stronger in presence of mixed objects than when the avatar was the only virtual content displayed. However, we could not conclude with confidence on the origins of the observed differences and their extent to larger populations. We discuss potential explanations as well as the implications and limits of our results.

Chapter 4. After investigating the SoE, we propose to explore how avatar embodiment in AR may be applied to enhance our perception of physical environments (RQ2). Using cognitive theories on affordance, we imagined a system letting users send a virtual version of themselves to act and acquire new information in the space ahead without having to enter it. This chapter details our investigation

of the interaction techniques required for such a system. We then present a proof-of-concept implementation combining three modes to control an avatar matching the user's body size and morphology (referred to as "self-avatar") in the third person perspective: Physical Control, Puppeteering, and Body Tracking. We finish this chapter by examining several ideas to complement this system.

Chapter 5. To continue, we present an exploratory study testing the approach of using self-avatars to improve real-world perception, presented in Chapter 4. Three experiments were run using two modes of the system presented in the previous chapter. The first experiment evaluated whether the participants made more precise judgments of their action capabilities ("affordances") when able to control their self-avatars in an environment set in front of them (RQ2). We found that the avatar effectively improved the user's accuracy and induced higher confidence when answering. The next two experiments assessed the limits of these effects over time by evaluating whether they could last even after removing the avatar from sight. Additionally, the third experiment also tested whether the judgment of affordance of virtual objects (a door, in this case) was impacted differently from that of physical objects. We could not find clear evidence that having manipulated the avatar influenced the user's performance in either experiment, in any of our conditions. We discuss these results and questions that they raise on user cognition.

Chapter 6. In the next part, we detail the ongoing prospective work conducted in collaboration with the CARE lab at Nara, Japan. We investigated how to exploit the time spent *before* the embodiment to reinforce BOIs and reduce their inter-subject variability (RQ3). To do so, we designed an experimental protocol where users were asked to imagine being their avatars and imitating them before entering its body. We evaluated the effects of such mental preparation in VR by measuring the Proteus Effects of a Hulk avatar. As the analysis of this experiment was still running during the redaction of this thesis, we only present preliminary results. From them, it seems that the users could partly benefit from a more immersive experience of their avatars. We additionally mention a secondary project in which we participated, led by Riku Otono of the CARE lab. This project explores the use of visual effects at different stages of the avatarization. The user study we conducted shows that using such transitions can extend the duration of the Proteus Effect and reinforce the user's sense of agency.

Chapter 7. Lastly, we conclude this thesis by taking a step back and looking at the different contributions presented in this manuscript to discuss their perspectives.

Research is to see what everybody else has seen, and to think what nobody else has thought.

Albert Szent-Györgyi

2

Related work

This chapter presents the general related work on body avatarization in AR and the sense of embodying avatars. It proposes a review of the literature taking into account both psychological and technological aspects. It does not aim to be an exhaustive collection, but to form a representative sample of the subject and to provide a rich description of it. Finally, it seeks to highlight the gray areas of AR embodiment research and to submit research proposals to clarify them. Related work more specific to each contribution can additionally be found in the respective chapters of this thesis.

Contents

2.1 *Applications of body avatarization in the AR literature*

Body avatarization in AR can be useful in a variety of applications. This section presents the use cases that were the most investigated in the AR literature for each part of the body avatarization continuum, introduced in Chapter 1.

2.1.1 *Body Accessorization*

Straightforwardly, one of the primary use cases for body accessorization in AR resides in retail. By overlaying products on the user's body, this type of augmentation can not only improve consumer experience (both in-store and online) but also remove potential discrepancies between products and perceived body sizes (Yim et al., 2019). Examples of AR try-on systems letting shoppers virtually fit clothes, jewelry, eye-wear, or shoes are abundant in both academia and industry (Eisert et al., 2007; Harrison et al., 2011; Hilsmann et al., 2009; Kleinberger et al., 2018; Saakes et al., 2016). More details on their experiential value can be found in existing surveys on that subject (Dacko, 2017; Parekh et al., 2020).

Body accessorization has also been used in entertainment, artistic and cultural activities. Its combination with AR provides users with a novel way to convey ideas by letting them change their attire dynamically. Among examples, the virtual make-up application MagicFace (Javornik et al., 2017) was deployed in a museum and in an opera house to let visitors observe how they would look with Egyptian pharaoh make-up. Its creators see a broader use for such applications: they could help make-up artists in their creation process or aid on-stage performers step into character without having to dress up. These ideas were explored by Bermano et al. (2017) who proposed a live facial make-up system designed for novel performances (see Figure 2.1). Treepong et al. (2018) also implemented a tool aiming to improve make-up creativity. In similar lines of work, AR dynamic dresses were proposed by Kleinberger et al. (2018) as a "new form of expression through clothing to reflect identity, personality, emotions and inner states of the wearer".

Figure 2.1: Example of body accessorization with virtual make-up. Figure extracted from the work of Bermano et al. (2017).

2.1.2 *Partial Avatarization*

Partial avatarization finds great relevance in the medical field. For instance, as in Figure 2.2, severed patients could use it to train themselves to use a virtual prosthesis before implanting a real one (Barbosa et al., 2012; Lamounier Jr et al., 2012). Research has shown that this could not only help to accelerate user adaptation but also reduce phantom limb pain and lower rejection rates through the development of long-lasting ownership illusions (Baskar et al., 2017; Nishino et al., 2017). Similarly, partial avatarization in AR could help to recover motor functions during rehabilitation after a stroke

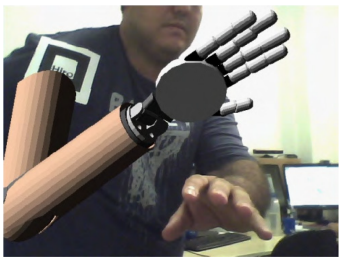


Figure 2.2: Example of partial avatarization. A user is augmented with a virtual prosthetic arm which he can see in AR. Figure extracted from the work of Lamounier Jr et al. (2012)

(Hoermann et al., 2012). Indeed, Heinrich et al. (2020) successfully used BOIs on paralyzed limbs to ease recovery by generating cognitive stimulation linked to body ownership, agency, and kinesthetic perception through visual feedback.

Instead of replacing or overlaying a human limb with a virtual one, partial avatarization can also be used to modify one's body structure and extend our possibilities of interaction. For example, Feuchtner et al. (2017) implemented an AR system providing users with an expandable virtual arm, allowing them to reach real objects over twice as far as they normally could (Figure 2.6) and interact with them through an IoT system. Lastly, T.-Y. Wang et al. (2019) showed that AR virtual hands could help people with hand tremors to type more efficiently. Their system tracked the user's hands and overlaid stabilized hand models over them in a slightly transparent way for the users to see the keyboard underneath. Their study illustrates how AR partial avatarization can be used to facilitate daily tasks for physically impaired or disabled people.

2.1.3 *Full Avatarization*

Full avatarization in AR extends what can be done with previous parts of the continuum: one could use it to embody an avatar with a different appearance, shape, and size, or to observe oneself in a given (real) situation through virtual out-of-body experiences. Examples of papers using full avatarization in the AR literature as in Figure 2.3 are extremely rare. Most of the research implementing such avatarization were user studies on perception and not application-oriented papers (J. v. Bommel, 2017; Nimcharoen et al., 2018; Rosa, J.-P. v. Bommel, et al., 2019; Wolf, Döllinger, et al., 2020).

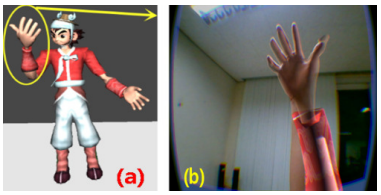


Figure 2.3: Example of full avatarization in first person perspective with a video see-through head-mounted display. The avatar is overlaid on the user's real body which is rendered partly transparent. Extracted from the paper of Noh et al. (2015).

Among the examples of applications we found, the study of Anderegg et al. (2018) shows how AR full-bodied avatars offer compelling possibilities of original game-plays. They were interested in novel ways to control full-bodied avatars in AR mobile games. The interaction metaphor they propose lets players control their avatars like a puppet with intuitive gestures. Hoang et al. (2018) and A. S. Johnson et al. (2013) illustrated how full avatarization can also be used in educational and health awareness settings. Both of them proposed spatial AR full avatarization systems exploring on-person anatomical displays for health education (muscles, blood vessels, etc.). Their results demonstrated a stronger level of "connectedness" and ownership with the projected content than with other projection supports, especially when the data was coherent with the participant's physiological state (e.g. synchronized heartbeat).

The lack of other examples in the literature is probably due to the difficulty of implementing full avatarization in AR (see Section 2.2.2). Nevertheless, one can imagine a large number of scenarios where full-bodied avatars could be useful. We discuss these potential use cases in Chapter 7.

2.2 *Implementing body avatarization in AR*

While the release of multiple frameworks (e.g. Unity ARFoundation, Windows MRTK) has considerably simplified the development of AR applications, setting up an SoE remains a truly complex task in AR. The purpose of this section is to constitute a guide to avatar embodiment in AR. We describe the core components of AR avatarization and break down the existing possibilities to implement strong ownership illusions in AR. Examples of applications will be given to illustrate the current uses of each *modus operandi*.

2.2.1 *The choice of user perspective*

To make body avatarization and ownership illusions possible, technical choices are to be made conscientiously as they could lead to very different user experiences. The first of these important decisions is from which perspective the avatar should be embodied.

In a First-person Perspective (1PP), the user will look through the eyes of their avatar and/or experience embodiment by looking down at their own body (e.g. Figure 2.4, left). This point of view can be combined with virtual mirrors that let users observe their body avatarization indirectly (e.g. Figure 2.4, center). In a Third-person Perspective (3PP), the avatar will be observed at an arbitrary location and result in an out-of-body experience (e.g. Figure 2.4, right). For example, instead of seeing a virtual arm where one's real arm would naturally be, one could see it floating ahead and use it to grab normally unreachable objects thanks to partial avatarization.

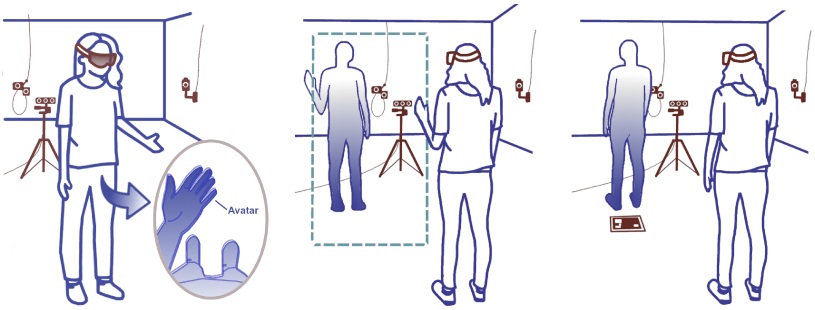


Figure 2.4: First- and third-person perspectives in an AR full avatarization setup. *Left:* The user wears an AR headset and looks down onto her fully avatarized body. *Center:* A virtual mirror reflects the user's body as an avatar. *Right:* The user observes her avatar instantiated ahead of a fiducial marker with a body pose identical to hers.

Two main factors may influence the choice of either perspective. The first is the application's purpose: does the user need to see specific parts of their body, like their face? If so, an egocentric point of view will only let users see virtual content at locations that they can directly see. Augmentations positioned on faces, necks, or backs will not be visible in this point of view and therefore limit certain avatarization applications unless combined with virtual mirrors. In the 3PP, users can observe their virtual avatarization in its entirety from different angles and standing viewpoints, providing them with an external awareness of their avatar.

The second factor to take into account is the strength of the ownership illusion that is sought. We are not aware of studies that have compared the 1PP and 3PP in AR. Nevertheless, the 1PP preserves the instinctive perception of self-location with respect to the real body and this seems essential to generate a strong SoE. Some VR studies support that the 3PP can generate equal ownership illusions as the 1PP (Debarba et al., 2015), but the major part of the community seems to agree that egocentric points of view will produce a stronger sense of ownership and self-location (Gorisse et al., 2017; Medeiros et al., 2018; Valeria Ivanova Petkova et al., 2011; Seinfeld et al., 2020; Slater, Spanlang, et al., 2010). This being said, 3PP can provide illusions strong enough to satisfy some AR embodiment applications (J. v. Bommel, 2017; Rosa, J.-P. v. Bommel, et al., 2019). This perspective also has the advantage of increasing awareness of the virtual posture and of its surroundings, (Debarba et al., 2015) and therefore justifies choosing it over the 1PP in some cases.

Ultimately, it seems recommended that applications seeking strong BOIs use the 1PP and combine it with virtual mirrors when applicable. More studies will be necessary to understand in which situations one perspective should be preferred over the other in AR.

2.2.2 *Displaying avatars in AR*

Avatar embodiment can be experienced through various AR displays with different modes of operation (e.g. head-mounted, handheld) and renderings (e.g. video, optical). In this section, we do not detail how these displays work or the advantage and issues that are specific to them[1]. Instead, we discuss how their characteristics make them more or less prevalent for embodiment experiences located at different parts of the body avatarization continuum.

[1] See the papers of Kruijff et al. (2010) and Zollmann et al. (2020) for a complete overview.

Non-immersive Video See-Through (VST) displays: they are monitor-based systems also known as Windows-on-the-world (WoW) displays. Smartphones and tablets are common examples (see Figure 2.5), but WoWs can also be larger screens. When implementing body avatarization, their main strength is that they can be combined with live image processing to erase the user's real body through in-painting techniques (Zollmann et al., 2020). This allows the partial and full avatarization of bodies with a morphology different from the user's that would otherwise be impossible (e.g. 1PP partial avatarization of a thinner arm). However, the screen sizes of WoW displays are often limited and will usually not let users see the whole virtual content at once. Most of the user's body consequently remains visible outside of the display during 1PP full-body avatarization. In previous work, such displays were mostly used to implement virtual mirrors (Eisert et al., 2007; Javornik et al., 2017; Qian et al., 2019; Treepong et al., 2018).

Figure 2.5: Example of partial avatarization with a tablet, borrowed from M. Kim et al. (2016). Other examples can be found among social network applications (e.g. SnapChat).

VST Head-mounted Display (HMD): this mostly refers to VR headsets that are equipped with external cameras (e.g. Oculus Quest 2, HTC Vive Pro). They seem better adapted to immersive 1PP embodiment experiences than non-immersive VST displays as they completely hide the direct vision of the user's real body. Combined with inpainting, 1PP full avatarization can be achieved at its best by completely overriding the real body's vision with the one of a full avatar, regardless of its morphology as done in Figure 2.6.

Figure 2.6: Example of partial avatarization with a VST HMD, borrowed from Feuchtner et al. (2017). *Left:* A virtual arm of normal length is embodied in first-person AR. *Right:* The user extends his virtual arm towards a normally unreachable (real) electric curtain and manipulates it.

Optical See-Through (OST) HMDs: unlike their VST counterpart, OST headsets provide an unmediated view of the real world (e.g. Microsoft Hololens, Magic Leap). On the one hand, this direct view "ensures that visual and proprioception information is synchronized" as described by Rolland et al. (2000). On the other hand, virtual content viewed with them is partly transparent and sometimes causes depth perception issues that may impact the embodiment experience, as witnessed by Gilbers (2017). Another issue with these headsets is that their usually narrow fields of view can break the continuity of ownership illusions as parts of the avatar might get cropped by the optical system and let the user see their real body in the outer areas of the display.

Projection-based systems: this refers to spatial AR and CAVE-like systems. Like in OST, the opacity of the content they render heavily depends on light output capacity and ambient lighting. Using it for partial and full avatarization is more complicated when the embodied limb or avatar is not matching the user's body. Stereoscopic CAVE systems provide more latitude, but removing the visibility of the user's body parts presents the same problems and may require ingenuity (as in Figure 2.7). Despite their restrictions, projection-based systems can nevertheless suit the needs of particular embodiment applications. For instance, A. S. Johnson et al. (2013) used SAR to allow users to interact with their own internal makeup projected onto their body (i.e. muscle tissues, blood vessels, etc.). SAR was also previously used for 3PP embodiment of differently shaped bodies: Gervais et al. (2016), for example, designed a small plastic figurine representing the user's inner state as a means of introspection. When embodying it, user physiological data is recorded and projected onto it.

Figure 2.7: Example of partial avatarization with a CAVE system, borrowed from Slater (2008). The subject sees the virtual arm as projecting out of his right shoulder, while his own arm is hidden by a panel.

OST "fishtank" setups: these systems use semi-transparent mirrors to reflect AR content displayed on a stereoscopic screen (see Figure 2.8). They can be used to overlay 3D content on and around the user inside of a dedicated tracked volume under the display's mirror (Genay et al., 2021b; Hachet et al., 2011; Mercier-Ganady et al., 2014; Shibuya et al., 2018; T.-Y. Wang et al., 2019). In general, this volume is small and will only allow body accessorization and partial avatarization (e.g. Shibuya et al. (2018)). OST fishtanks present the same drawbacks as projection-based systems.

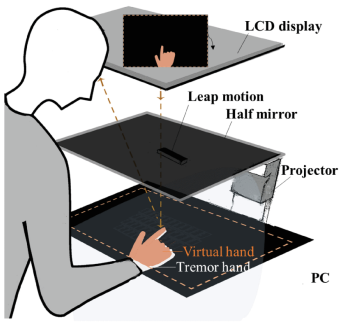


Figure 2.8: Example of OST fishtank display. Illustration taken from the work of K. Wang et al. (2017).

Overall, current trends seem to favor HMDs as the primary medium for AR display. HMDs have the advantage to combine free-hand interaction with multiple built-in sensors that provide realistic integration of virtual content (Makhataeva et al., 2020). Ultimately, however, there is no one-size-fits-all solution for all embodiment applications. Each trade-off must be evaluated in light of particular goals and available technologies to decide which form of display is best.

2.2.3 *Controlling and animating avatars in AR*

A component that is essential to body avatarization is the control of the embodied avatar. While it is possible to achieve an SoE without any avatar animation (Botvinick et al., 1998), giving control of it will enable interaction and strengthen the BOI (Ma et al., 2015)[2]. Virtual avatars can be animated by users in many ways, including with tangible interactions (Fernandez-Baena et al., 2014), controllers (Griffin et al., 2019), puppeteering (Anderegg et al., 2018; Oshita et al., 2013) or motion tracking technologies. These methods are not specific to AR and their usage for body avatarization is the same as in VR applications. However, they come with their own challenges when used in real-world contexts for embodiment.

[2] Agency is related to "behavioral realism", reflected in the number of behaviors of a human that the avatar exhibits. For example, allowing users to raise their avatar's hand in a classroom to ask a question will help them embody it (Schultze, 2010).

To our knowledge, no study explored the influence of different animation techniques on the SoE. However, the sense of agency is known to be greatly enhanced by visuomotor congruence between real and virtual body movements, while inconsistencies between visual and motor information appear to reduce it (Farrer et al., 2008; Sanchez-Vives et al., 2010). Control through inputs that are not physically representative of the user's body (e.g. with gamepads) might therefore not play in favor of strong BOIs. For this reason, tracking technologies that provide a 1-to-1 mapping of user movements have been the preferred tool for avatar control (Spanlang et al., 2014).

Numerous motion-tracking technologies exist, ranging from open-source software to costly hardware-based solutions. Optical markerless tracking seems to be the most common tracking technique in AR research (Eisert et al., 2007; Javornik et al., 2017; Rosa, J.-P. v. Bommel, et al., 2019). It uses computer vision to record the position and rotation of the user's body joints in real time. Free toolkits and libraries can be used to detect such information, e.g. OpenPose (Cao et al., 2018) or Unity ARFoundation. Once collected, these feature points are matched to the avatar's corresponding body or face landmarks and the 3D model is animated into the same pose or expression as the user (see Figure 2.9). Inverse Kinematics (IK) solvers are usually used to deduce the user's posture out of this data. They calculate joint angles based on body constraints so that only natural poses will be taken by the avatar (S. Kim et al., 2018).

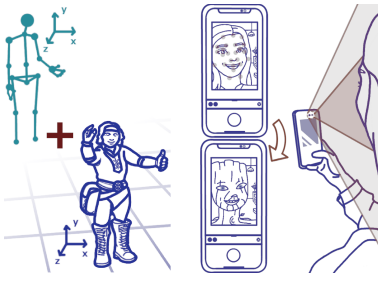


Figure 2.9: Animation process in AR. *Left:* A skeletal pose is acquired through user tracking and matched to animate the avatar into the estimated pose. *Right:* Face feature points are extracted to morph the avatar model for partial avatarization of the user's face. Here, it is visualized on a smartphone acting as a virtual mirror.

All of these processes make tracking and image generation inevitable sources of delay in AR systems, even with high refresh rates

(Jang et al., 2011). The effects of tracking errors on the SoE were not clarified in AR, but VR studies have found that certain forms of errors including latency, jitter, and noise can strongly affect the SoE (Waltemate, Senna, et al., 2016)[3]. According to Toothman et al. (2019), latency and noise (avatar vibration) appear to be particularly limiting the SoE at relatively low thresholds.

[3] The sense of agency is linked to the easiness of the avatar's manipulation. The less the user needs to think about how to control it, the more they will disregard the virtuality of their body (Bainbridge, 2004).

Noise and jitter are errors linked to the tracking signal which quality may vary importantly with the real environment's conditions (e.g. lighting). Providing robustness for these errors is still a hot topic of current research (Batmaz et al., 2020; Toothman et al., 2019). One solution to counter the latency issues of AR systems is to use VST displays. Indeed, a video view of the real world makes it possible to synchronize the overlay of virtual and real content by deliberately delaying video images to match the tracking and rendering latency (Zheng et al., 2014). On the other hand, the real-world view of projection-based and OST systems offers no latency and therefore makes it impossible to completely compensate entirely for delayed tracking: one can reduce misalignment between synthetic and real imagery through predictive tracking and selective update, but errors are still likely to be present, especially during rapid head pose changes (Zabels et al., 2021).

To conclude, motion-based tracking promises stronger illusions than its counterparts, but it comes with its share of technical difficulties in AR. Depending on the type of display used, it will introduce errors that can disrupt embodiment illusions and that may require effort to be handled. Latency is the most prominent source of registration error in existing AR systems (Lincoln et al., 2016). Because of the need to maintain alignment between real and virtual environments, avatarization applications in AR have to make an extra effort to preserve the user's sense of agency compared to VR.

2.2.4 *Integrating avatars to the real world*

Ultimately, it seems that AR embodiment systems rely on similar components as the core VR embodiment system depicted by Spanlang et al. (2014). The main difference they present resides in the type of displays used and in the replacement of the VR module with an AR one. However, this AR module has to accomplish an additional task that makes AR embodiment much more complex: it has to build spatial awareness of the physical environment to handle the proper integration of virtual content.

Several techniques exist to produce an accurate 3D registration of virtual avatarization in real space. The detection of fiducial makers, planar images, or 3D objects is a long-dating and simple way to position avatars at the desired location (Kato and Billinghurst, 1999), whether it be on the user's body (Lamounier Jr et al., 2012) or elsewhere (Rosa, J.-P. v. Bommel, et al., 2019). This method involves tracking markers registered beforehand in the AR system. Once the marker is recognized, the avatar is spawned in a location relative

to it and displayed. Such a process is fast to set up thanks to the availability of many free toolkits like ArUco (Muñoz-Salinas et al., 2014) or (Kato, Billinghurst, and Poupyrev, 2020), but alone, it does not provide the embodiment system with any understanding of the physical environment. Without such understanding, the relative position of real and virtual objects is ignored and the coherence of the scene cannot be ensured.

To provide occlusion between real and virtual objects, spatial awareness can be acquired through Simultaneous Localization and Mapping (SLAM). These techniques create an approximate model of the environment by using sensor data (mostly camera feed) to detect feature points, associate them with 3D space coordinates, and update them continuously (Durrant-Whyte et al., 2006; Reitmayr et al., 2010). Systems using SLAM do not require markers as the created model can be used to identify environment or body surfaces and place content onto them. Once detected, surfaces can also be virtually registered to stop virtual content from unnaturally entering walls and implement virtual occlusion, resulting in a stronger coherence (see Figure 2.10). On the downside, they require more computing power and tend to be less stable than marker-based systems as the feature points are built from sensor data that is updated continuously and that may contain errors.

Figure 2.10: An illustration of virtual object rendering both with and without occlusion rendering. The lack of occlusion makes the user's hand appear behind the virtual chest (top), whereas it should normally appear in front of it (bottom). Figure extracted from Du et al. (2016).

Examples of software capable of rendering AR content with spatial awareness are game engines (e.g. Unity, Unreal Engine). They are particularly fit for such tasks as they integrate avatar animation, state-of-the-art graphics, stereoscopic rendering, realistic physics, sound, memory management, and many AR development tools. Additionally, they support a fair number of commercialized tracking systems and allow cross-platform builds.

Finally, to draw a parallel with the core VR embodiment system described by Spanlang et al. (2014), we could define a core AR embodiment system composed of the following modules:

- *Tracking Module:* depending on the type of display (i.e. handheld vs. head-mounted), records the user's body and head movements or device's position in real-time. It allows controlling the avatar and adapting the AR scene's point of view. This tracking is used as input for the AR module.
- *AR Module:* processes the outputs of every other module and adapts the rendering of the AR scene according to them; takes charge of the integration of the avatar into the real world from the knowledge acquired through the tracking module or separate environment processing (e.g. SLAM); passes the resulting rendering to the display module.
- *Display Module:* hardware allowing the user to visualize their virtual avatar and the AR scene (see Section 2.2.2).

The AR module is also responsible for the management of the AR

content and usually handles the animation and interactions related to the avatar (i.e. pose estimation, IK, etc). In more advanced systems, it may also include lighting simulation, stereoscopic rendering, and audio. Additional modules can be added to this core system to provide multi-modal stimulation (e.g. haptic feedback) and measurement tools. Finally, alternative user inputs through controllers or voice can also be combined with the user tracking system to provide different embodiment experiences.

2.2.5 *Summary of technical challenges*

A variety of obstacles in adopting AR embodiment were mentioned in the previous sections. Most of them are linked to consistency issues of AR experiences. Dealing with these challenges includes:

- Providing correct registration and alignment of the virtual body with the real one during 1PP embodiment.
- Compensating for latency causing discrepancies between the user and the avatarization movements.
- Dealing with real environmental conditions that can impact the avatarization's visibility (e.g. lighting when using OST displays).
- Removing visible user body parts when implementing a virtual embodiment of differently structured bodies (e.g. thinner limbs).
- Ensuring that the virtual avatar responds realistically enough to the physical world (e.g. collide with walls, displays shadows).
- Providing a field of view sufficiently large to witness the avatarization as intended.

Dealing with these challenges can be more demanding or not apply depending on the targeted experience[4]. Achieving further realistic experiences of avatarization will require solving additional challenges such as letting users interact with real-world objects, rendering in high quality, or providing haptic feedback. In addition to these, research outlined other challenges linked to the general usage of AR, including visual fatigue, security issues, marker tracking reliability, and device costs (Akçayir et al., 2017; Masood et al., 2019).

However, it seems that these issues are relatively minor and should not prevent the adoption of AR for virtual embodiment. Most embodiment systems described before can already be implemented with state-of-the-art technology. Although some configurations are more challenging to achieve (e.g. 1PP full avatarization), new advances are anticipated to overcome current usability and technological issues such as low resolution, limited fields of view, cost, and low computing power.

[4] For example, simulating autoscopic experiences where one sees a duplicate of oneself does not require removing the visibility of the real body. We provide an example of this in Chapter 4.

2.3 Measuring the sense of embodiment in AR

As of today, measuring a psychological phenomenon as complex as the SoE remains an open debate. Very few publications share the same methods and choosing the most appropriate one is not always a simple task. The tools used in AR to measure the SoE are mostly derived from those previously used in VR, which themselves have evolved greatly over time.

2.3.1 Subjective measures

In the original BOI experiment, Botvinick et al. (1998) chose to design a two-part questionnaire to assess the SoE. The first part asked for an open-ended description of participants' feelings, while the second part consisted of nine questions to be answered on a Likert scale (Likert, 1932). Since then, numerous studies have been re-purposing this questionnaire, sometimes deleting the descriptive part or keeping this part only. These modifications have made the outcomes of studies particularly difficult to compare and validate.

In face of this, researchers have made efforts to unify embodiment questionnaires through systematic analysis. Longo et al. (2008) concluded that Likert scales are valid tools since it is possible to identify distinctive patterns in the subject's answers, but they should be supported by an objective measure that provides a complementary validation. In their view, the use of questionnaires alone is too subjective to extract an accurate measure of the SoE[5].

With more hindsight, research looked to establish standard questionnaires to improve comparability while accounting for these challenges. Roth, Lugrin, et al. (2017) identified a set of items through a principal component analysis of the literature and produced a questionnaire containing 12 items in 2020 (Roth and Latoschik, 2020). In the meantime, Gonzalez-Franco et al. (2018) proposed their own questionnaire built on exploratory factor analysis. It was updated in 2021 to keep a total of 16 items applicable to all scenarios, including those of AR (Peck and Gonzalez-Franco, 2021). Both questionnaires were tested and validated by several studies (4 and 9, respectively) and then by statistical reliability analyses (calculation of Cronbach's alpha coefficient).

The generic nature of these two questionnaires makes them appropriate for the evaluation of any type of AR avatarization. Additionally, their extended testing makes them good candidates to become the new standards. However, their construction and trial were mostly based on BOI and VR studies. It would be interesting to make a comparative study evaluating their efficiency in various contexts (including AR). Regardless, using one of these two validated questionnaires in AR will enable future comparison with experiments in VR. Using them in future AR research hence seems judicious, but it is recommended to implement control conditions to empirically test for potential bias linked to AR.

[5] Our relationship with our bodies is not something that we usually consider. Many people who have never lived or heard of autoscopic experiences may find questions like "I felt out of my body" unrelatable. Slater raised similar issues regarding the measurement of presence in VR with questionnaires (Slater, 2004).

2.3.2 *Objective measures*

Most objective measures used in VR can also be applied in AR. One that is commonly found in both is that of proprioceptive drifts, characterized by the change of a perceived body or limb's location towards a fake embodied one Botvinick et al. (1998). This measure can be obtained by subtracting the user's estimated body or limb position from its actual position (Gilbers, 2017; Škola et al., 2016). However, Rohde et al. (2011) and Holle et al. (2011) highlighted the risks of using this measure alone and the necessity of combining it with complementary ones as they showed that proprioceptive drifts could also occur in the absence of an SoE.

Bodily signals can be used in AR as indicators of BOIs too. For example, Skin Conductance Response (SCR) and heart rate deceleration (Rosa, Veltkamp, et al., 2019; Slater, Spanlang, et al., 2010) were previously used to indicate successful illusions, as in Figure 2.11. BOIs have also been correlated with decreases in user body temperature (Moseley et al., 2008) and variations of muscle electrical activity (Slater, Pérez Marcos, et al., 2008). A complete compendium of methods and technologies allowing us to obtain them can be found in the paper of Spanlang et al. (2014).

Lastly, brain activity can provide information on the neural processes involved. This activity can be analyzed through electroencephalography (EEG) recordings, as done in the right part of Figure 2.12. However, EEG recording has several drawbacks: on top of being time-consuming to set up, most non-intrusive EEG systems are very sensitive to muscle activity and electrical noise in the environment, limiting the embodiment to an immobile experience with very few possibilities of interaction. Further details on EEG setup and signal processing for the study of BOIs can also be found in the article of Spanlang et al. (2014).

While partial and full avatarization can be evaluated straightforwardly with any of these methods, the measurement of AR body accessorization experiences raises several questions: how does it relate in any way to embodiment experiences, and how can one quantify it? According to the survey of Schettler et al. (2019), objects can be experienced as being "a part or extension of one's body". To support this idea, they relate an experiment where researchers could observe and measure BOIs toward a table (Armel et al., 2003). The illusion was induced by stroking synchronously a table and the subject's hand after having applied a Band-Aid on both of them. Embodiment effects were measured with a questionnaire and SCR that showed that participants somehow felt that the table was their hand. Indeed, the questionnaire answers were corroborated by a peak in SCR upon partly pulling off the table's Band-Aid, leading users to expect pain [6]. These results suggest that body accessorization (whether in AR or not) does have an influence on body ownership and that this influence should be measurable with the same methods as in other avatarization experiences.

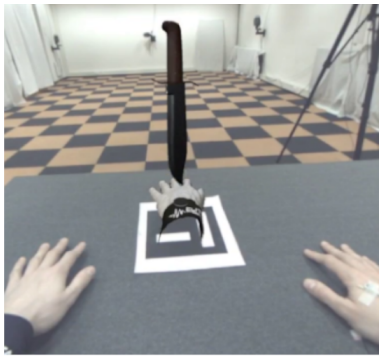


Figure 2.11: When the avatar is placed under a virtual threat (e.g. a sharp knife suddenly falls in this figure), studies have shown that participants will attempt to avoid virtual danger in the case where the illusion is successful. Sensors placed on the participant's right hand allow for the recording of their physiological response at the moment of the threat. Figure extracted from (Rosa, Veltkamp, et al., 2019).

Figure 2.12: EEG recording during visuotactile stimulation in a VHI in AR. The observation of increased activity in EEG bands can be used to detect the effects of BOIs. Figure extracted from (Škola et al., 2016).

[6] Neuropsychological research further supports the idea that inanimate objects can influence the body image and schema: Aglioti et al. (1996) described a case where a patient with a brain lesion had no sense of ownership over their arm but could recognize it again after removing their wedding ring.

2.4 Inducing the sense of embodiment in AR

Generally, the strategy used by researchers to induce the SoE in AR experiments is the same as in VR or in reality. In this section, we describe the methods classically used in literature to induce and maximize the embodiment senses in AR. The methods described here apply to body avatarizations from all parts of the continuum.

2.4.1 Visuotactile stimulation in AR

In the early days of research on the SoE, embodiment illusions were mostly induced through synchronous visuotactile (VT) stimulation. The original BOI experiment consisted in stroking a fake hand with a brush simultaneously with the participant's real hand, hidden from sight. Valeria I. Petkova et al. (2008) generalized this experiment to the whole body by providing the subjects with the point of view of a plastic human mannequin. While touching both bodies identically, they managed to create the illusion of embodying the mannequin. After testing VT stimulation on physical objects, the following question then arose: can such stimulation also be used to make one feel embodiment towards a 3D avatar? If so, how can this feeling be maximized?

To begin exploring this subject, Raz et al. (2008) designed a variant of the BOI called the Virtual Hand Illusion (VHI). This variant implements partial avatarization by replacing the rubber hand with a 3D virtual hand. Whether in AR or VR, the comparison made by Škola et al. (2016) showed that VT stimulation could evoke an SoE towards a virtual hand very close to the one evoked by a physical copy of this hand[7]. The VHI experiment was also generalized to the whole body to see whether VT stimulation could generate BOIs during full avatarization too. A large number of implementations were developed in VR, showing multiple times the effectiveness of VT stimulation with several types of avatars (Kilteni, Maselli, et al., 2015; Teräs, 2015). We could not find examples of full BOIs of avatars with VT stimulation in AR research. Research is lacking to verify that similar results would be obtained.

[7] In spatial AR, however, IJsselsteijn et al. (2006) obtained less conclusive results when projecting the virtual hand on a flat tabletop. The authors explained that the physical stroking of the flat hand projection rather than onto the rubber hand itself created inconsistency, disturbing the SoE. This problem being unnoticed in VR, their findings reveal the importance of graphic coherence AR.

2.4.2 Visuomotor stimulation in AR

Like VT stimulation, visuomotor (VM) stimulation can be used to induce embodiment illusions (Ehrsson, 2012; Kokkinara et al., 2014; Tsakiris, Prabhu, et al., 2006). It simply consists in replacing the tactile stimulation with the ability to move the virtual body. As with VT stimulation, synchronization of real and virtual movements was shown necessary to create embodiment illusions (Gilbers, 2017) VM stimulation is by far the most common technique to induce an SoE in AR research (J. v. Bommel, 2017; Rosa, J.-P. v. Bommel, et al., 2019; Shibuya et al., 2018; Wolf, Döllinger, et al., 2020). Compared to VT stimulation, VM stimulation is generally cheaper, easier to set up,

and easier to automatize (see Section 2.2.3). Automated VT stimulation in AR requires integrating haptic devices that can be more constraining or costly, especially when the user ought to physically move around and interact (Jeon et al., 2009).

On another note, the experiment of Slater, Spanlang, et al. (2010) showed that it is possible to exhibit a strong SoE in VR without having to stare at the embodied avatar. The authors explain that the sense of agency coupled with the first-person vision of a virtual world rich in realistic detail led to proprioception strong enough to suppress the need to look at one's body to feel like embodying it. Thus, the theoretical question of whether the absence of a direct view of the avatar can provide identical results in AR can be raised.

2.5 *Discussion*

The previous sections described experiments where the SoE could be evoked with different levels of body avatarization, embodiment perspectives, display types, and sensory stimulation. To some extent, previous work suggests that the SoE can be elicited independently of the technological context as long as the appropriate stimuli and induction conditions are met.

That being said, it is still not clear whether equivalent real, virtual, and hybrid environments will generate BOIs of the same strengths, nor that the SoE will be affected by the same factors in the same ways. The comparison made by Škola et al. (2016) led them to conclude that AR, VR, and real-world VT stimulation induce similar embodiment experiences, but their results also revealed some (small but real) differences between them[8]. Unlike in the real world or in VR where visuals are homogeneous, AR displays a graphical contrast that allows discrimination of the virtual from the real. As the SoE relies on the coherence of visual and proprioceptive signals (Kilteni, Maselli, et al., 2015), it seems likely that AR should provide a different perception of virtual content — and therefore of AR avatarization, depending on how well it is integrated. We follow up on this questioning in **Chapter 3**.

[8] Their subjective results notably showed a significant difference in the sense of ownership of a fake hand between the AR and real-world conditions, but not between AR and VR.

Secondly, this review shows that little research has explored interactions with one's avatar, whether it be in the 1PP or 3PP. Unlike the 1PP which imposes the use of body tracking (since the avatar must appear on the user at all times), the 3PP allows greater freedom in the ways to interact with one's avatar. For example, the 3PP makes it possible to control it with a joystick or to manipulate it at arm's length like a puppet. Although the lack of synchronization with the user's gestures may reduce the SoE in such cases, there are situations where a 1-to-1 body mapping is not possible (or wanted) and alternative modes of control become necessary. One striking application where different interactions are relevant is the use of 3PP avatars to explore and interact with surrounding real environments. We propose to discuss the design of systems enabling such exploration and contribute with lessons learned from the implementation of a proof-

of-concept in **Chapter 4**. Whereas previous research broadly studied the use of 3PP avatars to increase body awareness, very few have looked at how they may also increase spatial awareness. We propose to investigate this subject through a user study whose results are described in **Chapter 5**.

Lastly, it seems that the current state of technology does not allow the creation of completely perfect embodiment illusions, either in AR or VR. Consequently, the subjects need to be open to accepting the illusion for an SoE to be induced. However, as raised by recent research, the differences between individuals and their personalities lead them to be more or less willing (or capable) of suspending their disbelief, or even completely rejecting their virtual bodies (Schwind et al., 2017). This makes embodiment experiences very variable from one subject to another and has important impacts on the analysis of experimental results. Strategies to compensate for this variability were rarely investigated. Addressing this problem is crucial to allow avatars to be accepted and used in concrete application scenarios. One promising lead is to influence the user into accepting their avatar before its embodiment. In **Chapter 6**, we explore this track as well as the use of graphical transitions to smooth the end-to-end embodiment experience.

The most exciting phrase to hear in science, the one that heralds new discoveries, is not 'Eureka!' but 'That's funny...'

Isaac Asimov

3

Investigating how AR content influences the sense of embodiment

This chapter describes our work investigating the sense of embodiment of avatars embodied in AR with an OST display. We report a user study testing whether the content of mixed environments could impact embodiment illusions in the first person during partial avatarization with robot hands.

Contents

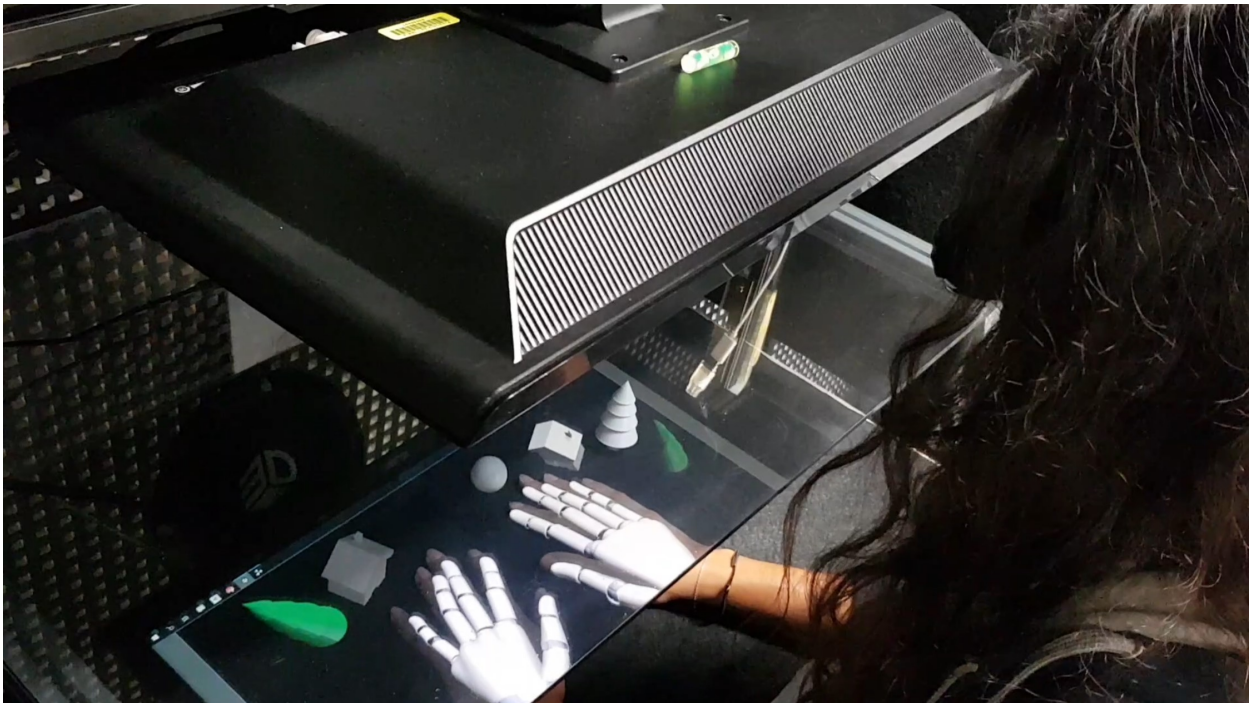


Figure 3.1: The experimental setup of our user study. A user is equipped with tracked shutter glasses and looks at her virtual hands under a semi-transparent mirror. The virtual hands are a reflection of the 3D screen hung above. The rendered perspective of the holograms is adapted to her point of view in real time. Part of the objects she sees is virtual (the green-colored ones).

Whereas embodiment illusions were widely studied in VR, little is known about their mechanisms in AR. The perception of AR is considerably different and it cannot be assumed that virtual avatarization is experienced the same way through it as in VR. In particular, it is unclear whether the contrast that makes virtual and real content distinguishable could impact the plausibility of the illusion. This contrast is even further amplified when using OST devices whose virtual rendering is partially transparent.

In this contribution, we looked into whether the consequent disruption of graphical coherence could cause the embodiment of avatars to be experienced differently in OST AR. Previous research showed that the transparency of avatars and the visibility of the user's real body were not necessarily prohibiting embodiment illusions (Rosa, J.-P. v. Bommel, et al., 2019; T.-Y. Wang et al., 2019). However, it remains unclear whether the mix of real and virtual content is impactful and to what extent. Investigating this seems of great value as it could mean that the context of the embodiment has to be taken into account during the analysis of future studies. Conversely, if mixing real and virtual content does not affect the SoE, then it could mean that results obtained in VR could be exploited for OST AR applications.

The following sections report the experiment we ran to investigate this subject. In a nutshell, our results provide original insights regarding the perception of partial avatarization in OST AR environments. We could notably observe significant differences in the body ownership induced by the various environmental conditions, suggesting that the content of the environment, or its "level of virtuality", could be a new factor of influence of the SoE in OST AR.

3.1 Related Work

Previous research identified many factors of influence of BOIs, including avatar appearance, embodiment perspective, sensory feedback, personality, and avatar control (Fribourg et al., 2020). Whereas avatar-related factors were intensively studied, very few papers have looked at whether the display and environmental context could play a part too. This section details previous work on this subject.

Body visibility and transparency

Several aspects of OST AR provide good reason to believe that embodiment experiences induced through it could evolve differently. First, the user's real body remains visible to some extent in AR. Feuchtner et al. (2017) showed that this visibility could indeed significantly hinder the strength of the BOIs in VST AR. However, other work also testified of strong BOIs even when leaving the entire real body in sight during 3PP avatarization (Rosa, J.-P. v. Bommel, et al., 2019; Wolf, Döllinger, et al., 2020). Therefore, it seems that its impact remains is limited when efficient stimulation is applied.

In addition to the real body's visibility, the transparency of OST displays also has to be taken into account. The avatar can appear less visible than the user's real body when the lighting conditions are unfavorable. K. Wang et al. (2017) attempted to clarify how much of a handicap this transparency can be by measuring the strength of VHIs displayed with different virtual:real light intensity ratios in 1PP AR (i.e. 0:1, 0.25:0.75, 0.5:0.5, 0.75:0.25, and 1:0). Their results show that ratios of 0.75:0.25 and 1:0 did not produce significantly different senses of ownership and agency towards the virtual hands. Therefore, it seems that partial transparency may not be a major impediment of virtual embodiment as long as the avatarization remains more visible than the user's real hands.

Graphical contrast

Lastly, VR and AR provide quite distinct visual feedback and it seems legitimate to question if the contrast between human vision and synthetic graphics could have an impact on the perceived plausibility of BOIs. Škola et al. (2016) looked at the differences between the traditional BOI and the VHI in both VR and AR. They found that AR and VR produced similar illusions, but weaker than when using

a genuine rubber hand. IJsselsteijn et al. (2006) also examined the same illusion in mediated and unmediated visualization. In their experiment, they replaced the rubber hand with a fake hand projected onto a tabletop. They found that applying VT stimulation with a real brush onto the projection caused a visual conflict between the real and 2D projected material. This conflict impeded the embodiment of the virtual hand, regarded as physically incoherent. Their results hence suggest that a proper 3D registration is necessary to induce embodiment illusions.

A more recent study conducted by Wolf, Döllinger, et al. (2020) investigated the differences between body weight perception in head-mounted VR and AR VST displays. During their experiment, subjects saw a photorealistic generic avatar through an interactive virtual mirror that offered VM feedback. The physical world viewed in the VST condition was faithfully reproduced in the VR condition. Unlike what the authors were expecting, participants did not report substantially different senses of ownership and presence in AR and in VR. They speculate that the VST video resolution and VR graphics might have been too comparable for users to experience a difference. To verify whether an unmediated vision of the AR condition with OST could have changed these results, Wolf, Fiedler, et al. (2022) ran a subsequent study with the same apparatus[1]. Unlike what they expected, they could find no significant difference between their OST condition and the VST one of the previous experiments. However, their descriptive results and post hoc power analysis are not convincing enough to rule out the influence of the used display. The authors conclude that further research is needed to confirm their findings.

[1] The results of their subsequent study comparing OST to VST were published after the work described in this chapter.

We found no experiment that evaluated BOIs while users interacted with mixed environments. The papers described above did not involve interactions other content than with their avatar. Additionally, they all provide their users with a third-person perspective of the avatarization. Perhaps manipulating real objects with a virtual body could result in a different perception of the avatarization, but also of the surrounding environment. Although not mandatory to induce the SoE, the control of avatars is likely to be a desired feature of applications involving embodiment. Extending previous research to avatars embodied in the first person is also of interest to many AR applications.

3.2 User experiment

The experiment we conducted let participants embody virtual robot hands within three environmental conditions (see Figure 3.2).

- **Condition 1** (VIRTUAL) The environment contained only virtual items, laid in front of the participant's hands.
- **Condition 2** (MIXED) Some of the virtual items (half of them) were replaced by their real version.
- **Condition 3** (REAL) Except for the avatar hands, all content viewed by the participant was real.

One must note that all of these conditions took part in a real environment as we were interested in comparing OST experiences of embodiment. Therefore, the term "VIRTUAL" for condition 1 does not refer to a VR condition, but to the virtual objects and hands that are featured in the (real) scene. In each condition, the user saw these hands overlaid on their own hands and performed a VM task aiming to induce an SoE over them.

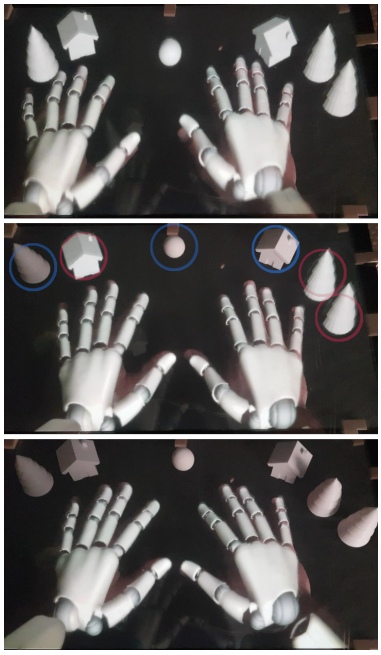


Figure 3.2: The experimental conditions from the participant's point of view. *Left* The "VIRTUAL" condition where all the objects were virtual. *Center* The "MIXED" condition where there were both real and virtual objects mixed. *Right* The "REAL" condition where all objects were real. The 3D rendering allowed virtual objects to appear having the same 3D volume as their real counterparts.

3.2.1 Participants

Twenty-six participants from age 21 to 61 ($\overline{m} = 30.8, SD = 12.2$ with 15 men and 11 women) completed our experiment. Most of them were colleagues from the research center and students from the university campus. They all volunteered and most were curious about AR technologies. No compensation or course credits were issued, and all participants were unaware of the purpose of the experiment.

Among them, 11 reported having no prior experience with AR or VR, 10 had little prior experience, and 5 had used these technologies many times. All had normal or corrected-to-normal vision and those with glasses could keep them on if not causing discomfort (one participant removed her glasses). Lastly, this study was approved by a local committee of ethics.

3.2.2 Apparatus

To control the visualization of both real and virtual contents in a seamless space, we built a setup inspired by previous VR fish-tank installations (Hachet et al., 2011; Hilliges et al., 2012). Our experimental setup is composed of a 3D display (NVIDIA 3D Vision, 120 Hz), a semi-transparent mirror, and a board for object placement. When equipped with shutter glasses, it allows subjects to observe semi-transparent "holograms" in the interactive volume located between the board and the mirror. These holograms are in fact the reflection of the 3D images displayed on the stereoscopic screen suspended above (see Figure 3.3).

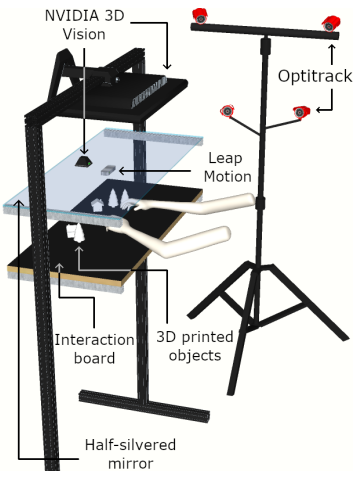


Figure 3.3: The experimental setup in front of which the participants stood. Reflective markers were attached to the shutter glasses worn by the participants for head tracking.

The experiment was developed using Unity 3D 2019.2. When positioning their forearms in the interaction volume, participants could see the forearms of an avatar co-located with their own and following their movements (see Figure 3.4). We used robot hand models extracted from the 2.3.1 version of the Leap Motion Unity SDK which was also used to track and animate the avatar's hands.

Secondly, to ensure a physically correct parallax, we implemented head tracking with Optitrack cameras. Indeed, the rendering of the 3D objects had to take into account the user's point of view in real-time for volumetric cues to be displayed correctly (e.g. the user should see the side of an object when moving their head to the side). This adaptation was achieved through a custom shader implementing 3D anamorphic projections that distorts the rendering according to head tracking inputs.

The experimental device was installed in a closed room where reflective objects were masked to reduce infrared interference. To reinforce the illusion of the presence of the 3D content, we simulated the lighting conditions of the room, illuminated with a single projector to simplify the simulation. As the semi-transparent mirror reflects only 70% of the incident light, the projector lighting makes the user's real hands slightly visible underneath the mirror, resulting in a rendering close to the one proposed in commercialized OST headsets.

The environment observed during the experiment consisted of empty space with six virtual and/or real miniature objects (houses, trees, balls), modeled and printed in 3D (see Figure 3.4). Depending on the experimental condition, a number of these real or virtual objects were positioned on the interaction board, in front of where the participant would place their hands. As the system did not provide haptic feedback for the virtual objects, the participants were instructed not to touch any of the objects in order not to introduce a difference in perception between the real and virtual objects. Once the experiment started, instructions were given to the participants by playing an audio recording.

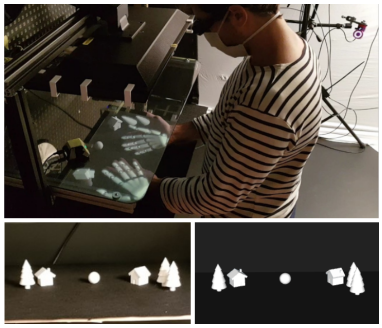


Figure 3.4: *Top* A participant performing the task of turning his hands over during the MIXED condition. *Bottom* Illustration of the object configuration that was seen by the user. *Left* Real 3D-printed version. *Right* their virtual counterparts.

3.2.3 *Procedure*

Participants started by completing a consent form containing the written instructions for the experiment. After verbal explanations and video demonstration of the task, they were asked to fill out a questionnaire assessing their Locus of Control (LoC) for further analysis (cf. Section 3.2.5). The participants were then invited to stand in front of the experimental setup and place their hands flat on the interaction board to proceed with a short calibration, lasting approximately 2 min. This calibration aimed to align the virtual hands with the real ones by moving them (if needed) in the direction expressed by the participant. While they were acclimating to the system, the experimenter invited the participant to ask questions and requested them to focus on their virtual hands.

The experiment was divided into three blocks that ran sequentially (one per condition), each displaying different amounts of real/virtual content. Depending on the block, miniature houses, trees, and a small ball appeared either really or virtually (or both) laid ahead of the participant's hands on the board. Each block differed only in the nature of these objects, which preserved the same appearance and position. The experiment then involved a VM task consisting in successively reproducing gestures dictated by audio instructions, which were presented beforehand. More specifically, participants had to (i) drum with their fingers, (ii) flip their hands, (iii) adduct/abduct their fingers, and (iv) position their right/left hand in front of a given item. These gestures, repeated 3 times each, had to be executed during a time delimited by a start "beep" and a stop "beep" for a total duration of about 4 minutes.

At the end of the audio recording, participants left the experimental setup to complete an SoE questionnaire adapted from the 7-point scale questionnaire of Gonzalez-Franco et al. (2018) (cf. Section 3.2.5). They were then invited to proceed to the next experimental block and repeat the same steps. The same avatar, task, and object configurations were maintained, but some virtual objects were replaced by their real version (or vice versa) depending on the condition tested. After going through all blocks and answering their respective questionnaires, participants were asked to fill out one last questionnaire assessing their post-experiment perception of the avatar and different environments. The average total time per participant, including instructions, questionnaires, experiment, breaks, and debriefing, was 1 h 30 min.

3.2.4 *Study design*

Previous research indicates that there are considerable differences between individuals in their ability to experience BOIs (Dewez et al., 2019; Richard et al., 2022). As between-subject designs are very sensitive to inter-individual differences, we chose to use a within-subject design to monitor these variations and increase the sensitivity of our experiment. The conditions were preceded by an acclimation phase of about one minute and counterbalanced with a Latin Square.

The choice of the avatar's robotic appearance was based on the recommendations of Gilbers (2017) who found that human virtual hands embodied in AR provoked high criticism from participants toward their realism and likeness. Several studies have similarly observed that VHI experiences should provide gender-matching models (Schwind et al., 2017) and color-matching skins (Lira et al., 2017) to avoid user self-comparison with the embodied avatar (see Chapter 2). Our experiment being focused on the influence of mixing realities and not on the avatar's appearance, we did not want to provide personalized avatars that would potentially create inter-individual variations. Therefore, we chose a generic non-gendered robot model to avoid bias and increase result comparability.

For consistency, objects were arranged in the same way for all participants, and replaced objects were always the following: the sphere in the middle, the tree on the left, and the house on the right (see Figure 3.4). The experimenter created the condition arrangements during the completion of the embodiment questionnaires, without the participant witnessing this change. Unlike the virtual hands, one must note that the virtual objects were not superimposed over their real counterparts. We decided to proceed this way because it is the most common situation encountered in head-mounted OST experiences: holograms are displayed in vacant space while virtual hands (when provided) are texture overlaid on the user's hands.

Regarding the task, the drumming and adduction/abduction gestures were inspired by the experiment of Hoyet et al. (2016) which used them to induce SoE toward a six-fingered hand in VR. The third gesture used in their study (opening/clenching the fists) was replaced by the gesture of turning/flipping the hands because it caused fewer tracking artifacts. The fourth gesture is of our design: it aims to momentarily (5 seconds) shift the participant's attention to the displayed environment so that they could take it into account when evaluating how they felt about the avatar.

To limit potential habituation and practice effects, each of these four gestures was repeated in a random order, renewed at each condition. The duration of the task (4 min) was chosen so that the stimulation would be long enough to induce an SoE (Lloyd, 2007) and short enough not to cause weariness.

3.2.5 *Collected data*

Out of the 26 participants, the data of 4 was excluded from further analysis due to technical issues compromising the avatar's perception. For the rest, we analyzed the data from the questionnaires presented in the following subsections.

3.2.5.1 *Locus of control*

The LoC is a personality trait corresponding to the "degree to which people believe that they have control over the outcome of events in their lives as opposed to external forces beyond their control" (Rotter, 1966). This trait was shown by Dewez et al. (2019) to be correlated positively with the sense of ownership and agency. Its measurement hence provides an additional tool to get a sense of the variability in the obtained results. A translated version (Rossier et al., 2002) of the 24-item IPC scale of Lefcourt (1981) was administered before the experiment's start. This scale provides scores between 0 and 48 for three dimensions: "Internality", "Powerful Others", and "Chance".

3.2.5.2 *Embodiment questionnaire*

Despite a large number of studies on the SoE, there still is no gold standard to evaluate this sense. We chose the widely used 7-point Likert scale questionnaire proposed by Gonzalez-Franco et al. (2018), built with the most prevalent questions found in the literature. This questionnaire covers six areas of interest: body ownership, agency, tactile sensations, location, external appearance, and response to external stimuli. We removed the questions unrelated to our experimental settings (e.g. statements on mirrors, haptic feedback, non-collocated virtual bodies, etc.) to keep a total of 12 questions (see Table 3.1). Their answers can be computed into sub-scores by a group of interest and into a global embodiment score.

Table 3.1: The embodiment questionnaire, adapted from Gonzalez-Franco et al. (2018).
We report the median and the first and third quartiles for each item. The "Score" column gives the contribution of each item to the following scores: Ownership (O), Agency (A), Location (L), and Appearance (AP).

ID	Score	Questions	Median[Q1,Q3]		
			VIRTUAL	MIXED	REAL
Q1	O	"I felt as if the virtual hands were my hands."	$1[-0.5, 2]$	$1[-0.5, 2]$	$1[-1, 1]$
Q2	O	"It felt as if the virtual hands I saw were someone else's hands."	$-2[-2.75, -2]$	$-2[-3, -2]$	$-2.5[-3, -2]$
Q3	O	"It seemed as if I might have more than two hands."	$-2[-2.75, -2]$	$-2[-3, -2]$	$-2[-2, -2]$
Q4	A	"It felt like I could control the virtual hands as if they were my own hands."	$1.5[1, 2]$	$1[1, 2]$	$1[1, 2]$
Q5	A	"The movements of the virtual hands were caused by my movements."	$2[2, 3]$	$2[2, 3]$	$2[2, 3]$
Q6	A	"I felt as if the virtual hands were moving by themselves."	$-2[-2, 1]$	$-2[-2, -1]$	$-2[-2, -1]$
Q7	L	"I felt as if my hands were located where I saw the virtual hands."	$2[1, 2]$	$2[1, 2]$	$2[1, 2]$
Q8	L	"I felt out of my body."	$-2[-3, -0.25]$	$-2[-2, -1]$	$-2[-3, -1]$
Q9	AP	"It felt as if my (real) hands were turning into "avatar" hands."	$1[-1, 2]$	$1[-1, 2]$	$1[0, 2]$
Q10	AP	"At some point, it felt as if my real hands were starting to take on the posture or shape of the virtual hands that I saw."	$0[-2, 1]$	$1[-1, 1]$	$0.5[-2, 1]$
Q11	AP	"At some point, I felt that the virtual hands resembled my own (real) hands, in terms of shape, skin tone or other visual features."	$0[-2, 1]$	$-0.5[-1.75, 1]$	$1[-2, 1]$
Q12	AP	"I felt like I had different hands from when I came to the laboratory."	$-1[-2, 1]$	$-2[-2, 1]$	$0.5[-2, 1]$

3.2.5.3 *Environment perception*

Given our interest to explore the impacts of blending real and virtual content, we designed a post-experiment questionnaire to assess the perception of the environments we tested. These questions were tailored to address aspects we felt could influence the SoE, such as the feeling of the objects' presence, the realism of the observed content, the feeling of being "immersed", and the mental workload demanded by the task. The participants were asked to remember each condition and then to rate them on a 6-point Likert Scale, ranging from "strongly disagree" to "strongly agree".

We also added two open-ended questions in order to get feedback in the participant's own words. The first one asked the participants to describe their feelings regarding their virtual hands and if there were differences between the three conditions. Because user appreciation is important in the design of such experiences, the second question asked the participants to indicate if they had a preferred condition and to rank them from most favorite to least preferred. The full 8-item questionnaire is detailed in Table 3.2.

Table 3.2: Overview of the post-experiment questionnaire responses.
For each question, we report the median and the first and third quartiles. P3 and P5 do not apply to the VIRTUAL condition as they did not contain real objects. Answers up to P6 were coded from -3 to $+3$.

ID	**Questions**	**Median[Q1,Q3]**		
		VIRTUAL	MIXED	REAL
P1	"I felt like I was immersed in a virtual world."	$2[1, 2]$	$1[-1, 2]$	$1[-1, 1]$
P2	"I felt that the virtual hands were visually well integrated into the environment."	$2[1.25, 2.75]$	$1[1, 2]$	$1[1, 2]$
P3	"I observed a visual inconsistency between virtual and real content."		$1[-1.75, 1.75]$	$-1.5[-2, 1]$
P4	"The virtual content I observed seemed to be really present."	$2[1, 2]$	$1.5[-0.5, 2]$	$2[1, 2]$
P5	"The presence of real objects bothered me."		$-2[-3, 1]$	$2[1, 2]$
P6	"Performing the tasks demanded concentration."	$-2[-2, 1]$	$-1[-2, 1]$	$-2[-3, -1]$
P7	Describe how you felt about your hands during the experiment. Were there any differences between the three conditions? Please explain.			
P8	In the future, if you had to use a virtual body, which condition would you prefer? Please rank the conditions by order of preference and justify.			

3.3 *Results*

Each of the experimental interests surveyed in the embodiment questionnaire (i.e. ownership, agency, location, and appearance) were computed into separate scores as described in the original questionnaire (Gonzalez-Franco et al., 2018): the textual ordinal answers were first converted to numerical data ("strongly disagree" $\rightarrow -3$, "disagree" $\rightarrow -2$, etc.) and then summed by group of interest. Answers of control items were inverted (e.g. Q2 for Q1). The resulting scores were then aggregated and averaged to compute a global embodiment score ("total SoE") by following the original questionnaire's method. The individual contribution of the questions to each sub-score is detailed in Table 3.1, "Score" column.

The LoC was also calculated into three scores corresponding to the Internality, Powerful Others, and Chance dimensions of Levenson's scale (Lefcourt, 1981). As for the post-experiment questionnaire, the questions were analyzed one by one without resorting to scores. Their answers (on a 6-point Likert Scale) were converted to numerical data (same process as before) to perform a descriptive analysis. The results presented below were obtained with R.

3.3.1 *Locus of control*

In order to have an overview of our sample's LoC profiles, we produced descriptive statistics on the LoC scores. The medians for each dimension (i.e. Internality, Powerful Others, and Chance) were $M_I = 18, M_P = 21$ and $M_C = 19$ with standard deviations of $SD_I = 5.4, SD_P = 5.3$, and $SD_C = 4.9$. We performed a cluster analysis (K-Means) and found that participants could be categorized into 3 groups: one with relatively high Internality and Powerful Others scores ($n = 12$), one with relatively high Powerful Others and Chance scores ($n = 6$), and one with low scores for all dimensions ($n = 4$).

3.3.2 *Embodiment scores*

Generally speaking, all conditions appear to have successfully evoked a positive SoE. A summary of the descriptive analysis of each score is presented in Figure 3.5 and Table 3.3.

Table 3.3: Descriptive statistics of the embodiment scores. We report the median and the first (Q1) and third (Q3) quartiles for each item in the following format: Median[Q1,Q3].

	Ownership	**Agency**	**Location**	**Appearance**	**Total SoE**
VIRTUAL	1.83[1.17, 2.00]	2.00[1.00, 2.25]	1.50[0.63, 2.00]	-0.25[-0.75, 0.69]	1.46[0.81, 1.77]
MIXED	1.83[1.33, 2.00]	1.67[1.08, 2.00]	1.50[0.63, 2.00]	0.13[-0.50, 0.94]	1.26[1.06, 1.71]
REAL	1.33[0.75, 2.25]	1.67[1.00, 2.58]	2.00[1.13, 2.38]	0.00[-0.44, 0.94]	1.14[0.76, 1.84]

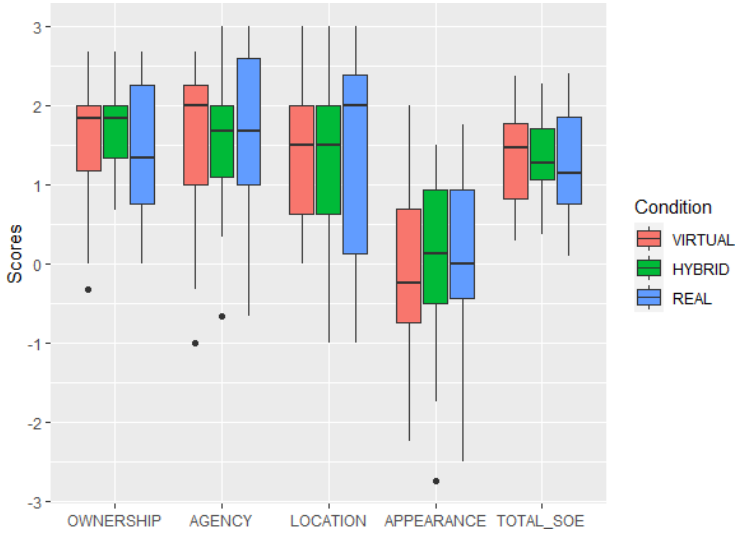


Figure 3.5: Boxplots representing the averaged embodiment scores for each condition. The scores all range from -3 to $+3$.

To identify differences between the SoE of each condition, we ran separate statistical tests on the SoE scores. Shapiro-Wilk and Mauchly's tests showed that none of them met the assumptions for an ANOVA. Therefore, we applied Friedman's tests as it takes into account the ordinal nature of Likert scales (Table 3.4). Only the ownership score came out as significantly different across the tested environments ($p < 0.05$) with a small effect size (Kendall's $W = 0.194$).

	χ^2	$p-value$		VIRTUAL	MIXED
Ownership	8.541	0.014*	**MIXED**	1.000	
			REAL	0.444	0.006**
Agency	0.585	0.747			
Location	1.099	0.577			
Appearance	0.469	0.791			
Total SoE	2.546	0.280			

Table 3.4: (Left) Friedman test results on the embodiment scores. Only ownership presented a significant difference. (Right) Wilcoxon signed-rank test p-values with Bonferroni correction, applied only on ownership ($*$: $p < 0.05$, $**$: $p < 0.01$).

Subsequently, we applied a posthoc test (Wilcoxon Signed-rank) with Bonferroni correction to determine pairwise differences. We found that this significant result was issued from the comparison between the MIXED and the REAL conditions ($p < 0.05$), other pairs being non-significantly different. The VIRTUAL (V) and MIXED (M) conditions had equal score medians ($M_V = M_M = 5.5$, $SD_V = 2.5$, $SD_M = 1.8$), superior to the one of the REAL (R) condition ($M_R = 4.0$, $SD_R = 2.3$). Further details are given in Table 3.4.

3.3.3 *Post-experiment evaluation*

To get a first idea of the post-experiment questionnaire results, we ran a descriptive analysis summarizing the main features of the data (see Table 3.2). The distribution of answers to questions P1 to P6 is illustrated in Figure 3.6.

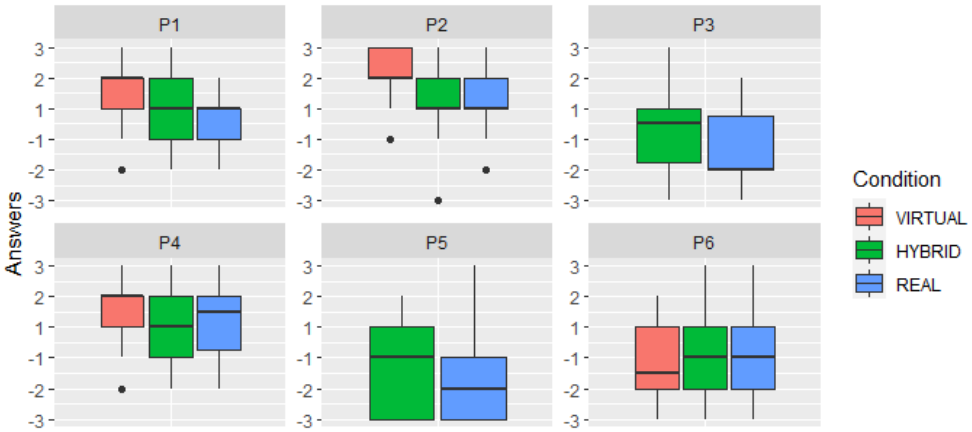


Figure 3.6: Boxplots of the post-experiment questionnaire answers, evaluated on a 6-point (forced) Likert scale. Answers were coded from -3 ("strongly disagree") to $+3$ ("strongly agree").

Next, we analyzed whether these six statements were significantly different across conditions. To do so, we ran Friedman tests on each of them (all questions were non-normally distributed) and found significant differences in questions P1 ($p < 0.001$) and P2 ($p < 0.01$) (see Table 3.5). The calculation of Kendall's W shows the effect of these tests to be moderate for P1 ($W = 0.319$) and small for P2 ($W = 0.288$).

	χ^2	$p-value$		VIRTUAL	MIXED
P1	14.033	$< 0.001^{***}$	**MIXED**	0.292	
			REAL	0.005^{**}	0.456
P2	12.689	0.002^{**}	**MIXED**	0.211	
			REAL	0.010^{**}	0.775

Table 3.5: Significant differences found in the post-experiment questionnaire results across tested environments. (Left) Friedman test results, (Right) pairwise Wilcoxon signed-rank test results with Bonferroni correction ($* : p < 0.05, ** : p < 0.01, *** : p < 0.001$).

Posthoc tests with Bonferroni correction revealed that the significant differences resulted from the VIRTUAL-REAL pair in both cases ($p < 0.01$ and $p < 0.05$ resp.). For most participants, medians seem to indicate that they scored the highest in the VIRTUAL condition for P1 and P2, followed by the MIXED and REAL conditions.

To identify a potential link between the perception of the environment and the SoE, we applied Spearman correlation tests between the SoE scores (all conditions together) and questions from P1 to P6. Positive correlations appeared with P1 (immersion), P2 (avatar integration), P4 (object presence), P5 (mix discomfort), and the embodiment scores detailed in Table 3.6. There was also a negative correlation between P6 (mental workload) and the total SoE score, but no significant difference was found between the conditions regarding this question.

	Ownership	Agency	Location	Appearance	Total SoE
P1	0.29*			0.37**	0.26*
P2	0.30*	0.26*		0.29*	0.24*
P3					
P4				0.25*	
P5	0.32*			0.40**	
P6					-0.27*

Table 3.6: Spearman correlations between the embodiment scores and the post-experiment questionnaire ($* : p < 0.05, ** : p < 0.01, *** : p < 0.001$).

Regarding the P7 open-ended question, several topics came out as prominent: the ownership/disconnection of the avatar, the evolution of this feeling, the system's quality, the separation/merge of virtual and real worlds, and the enjoyment of the experiment. Participants ordered their most preferred to least preferred conditions as follows:

(i) VIRTUAL > MIXED > REAL was the most frequent ranking (6 participants).

(ii) MIXED > VIRTUAL > REAL came next with VIRTUAL > REAL > MIXED (5 part. each).

(iii) Other rankings appeared 3 times or less.

(iv) One participant could not rank the conditions, being undecided.

Overall, VIRTUAL seemed to be the most preferred condition (half ranked it as their first choice) and REAL the least preferred one (half ranked it as their last choice), but we could not confirm this preference statistically (Friedman $\chi^2 = 3.7143$, p-value $= 0.1561$). We review these results in Section 3.5.

3.4 Additional Analyses

To put the previously presented results to the test, we conducted several post hoc analyses. Their results are presented in this section.

3.4.1 First trials analysis

Although it is of reduced power, the analysis of the first trials of each participant can be interesting to check if the results would have been the same in a between-subjects design. Indeed, the first trials can be grouped by condition to simulate independent measures. We applied such grouping and performed a Kruskal-Wallis test to compare the distributions of the three groups. No significant results were found for group sizes of 8, 7, and 7 for VIRTUAL, MIXED, and HYBRID respectively.

3.4.2 Updated embodiment questionnaire

As mentioned in Section 3.2.5.2, the 12-item SoE questionnaire we used was adapted from the questionnaire of Gonzalez-Franco et al. (2018). This questionnaire was updated after the completion of our study (Peck and Gonzalez-Franco, 2021): 9 items were removed from the original version and a new score computation method was proposed.

As the 16 items of this updated version are a subset of the previous questionnaire, we could run a second analysis with the revised version's methodology. The goal of this recalculation was to extend the validity of our results by making our analysis forward-compatible with papers that adopted the 2021 questionnaire. To do so, we removed the items that were not present in the new version and recomputed the scores following the newly described method. Using this method, we could not observe the significant difference that we had found between REAL and MIXED in terms of Ownership scores. We discuss this outcome in Section 3.5.7.

3.4.3 Per-item embodiment analysis

To understand better how the previously used items were involved in the significant results we had found, we performed a per-item analysis comparing each item of the embodiment questionnaire across the 3 conditions. Through this process, we aimed to identify which questions contributed to the significant difference in the Ownership score (Section 3.3.2).

The per-item analysis showed that Q1 had significantly different answers across the conditions (Friedman chi-squared $= 9$, p-value $= 0.011$). A post hoc test (Wilcoxon signed-rank with Bonferroni) revealed that this difference occurred between the conditions MIXED and REAL, matching with the Ownership score analysis. Other questions did not provide a significant difference.

3.5 Discussion

Whereas most of the previous research on BOIs was centered on the avatar itself, we contribute with novel insight into the play of environmental context on the SoE. Our goal was to explore how the blending of realities experienced in OST systems can modulate virtual embodiment sensations. The results we obtained reveal a potential difference in virtual body ownership linked to the amount of virtual/real content seen:

- Participants perceived the virtual hands significantly more as their own in the condition where both kinds of objects were mixed (MIXED) than in the one where only real objects (REAL) were in their field of view (Table 3.4).
- On the other hand, displaying unmixed virtual objects (VIRTUAL) created similar body ownership scores as in the MIXED condition (Figure 3.5).
- Meanwhile, the difference that could be expected between the VIRTUAL and REAL conditions did not appear (Table 3.4).

The other embodiment factors we evaluated (i.e. agency, self-location, appearance, and total SoE) did not emerge as significantly different. The fact that body ownership evolved separately is not without precedent. Indeed, previous studies have shown that SoE sub-components could be dissociated and elicited independently but that their co-occurrence would lead them to strengthen each other (Braun et al., 2014; Kalckert et al., 2012). On the other hand, the differences we observed between the MIXED and REAL conditions raise novel questions to which we attempt to provide an explanation.

3.5.1 Feelings of integration and immersion

As the same protocol, tracking system, and avatar were used for the three conditions, the origin of the ownership differences we witnessed would logically be related to the variations in the objects set up in the environment. The first avenue we explored therefore consisted in evaluating how each environment was perceived and in identifying potential differences in their cognition.

From the post-experiment questionnaire, the VIRTUAL condition appeared as significantly superior to REAL in terms of (i) feeling of immersion and (ii) feeling of "integration" of the virtual hands in the real environment (Figure 3.6, P2). More generally speaking, these two feelings were stronger when the virtual objects were present in larger quantities. This is reflected by the percentages of participants who agreed or strongly agreed to these statements: 59.1% and 77.3% did in VIRTUAL for P1 and P2 respectively, 45.5% did in MIXED (for both), against 22.7% and 31.8% in REAL.

Given these decreasing ratings, it seems as if replacing a part of the real items with virtual ones somewhat helped to provide immer-

sion and coherence to the avatar. We investigated a potential relationship with the SoE by applying Spearman correlation tests between the P1/P2 answers and each of the embodiment scores (all conditions taken together). P2 showed moderate positive correlations for all scores except location, whereas P1 did for all except location and agency (Table 3.6). In other words, regardless of the condition, participants who thought the virtual hands were well-integrated also tended to score higher on these specific SoE dimensions. Similarly, participants who felt more immersed also scored higher on these embodiment scores (and vice versa).

"Immersion" is a psychological state linked to the awareness of one's own environment and physical state (Slater, 2018). The entire experiment being visualized in OST, participants were certainly more aware of the real world than in previous BOI studies in VR. Nevertheless, the conditions appeared to be different enough for participants to notice a preference for VIRTUAL in P1 and P2. The fact that this condition had higher immersion ratings seems somehow logical as the virtual content occupied considerably more visual space in it. This visual occupation could have led participants to be more distracted from their real hands, and therefore to be more prone to develop body ownership in this condition. This would be in line with previous studies that found immersion to improve the SoE in VR (Waltemate, Gall, et al., 2018).

On the other hand, it remains unclear as to why the ownership scores of VIRTUAL and REAL did not emerge as significantly different when the P1 and P2 ratings were found to be significantly lower in REAL than in VIRTUAL. This is especially puzzling since MIXED and REAL did show a significant difference in their ownership scores, but not in P1 and P2 answers. In this regard, it should be kept in mind that the correlations mentioned above are only moderate and that other factors of influence could be at play.

3.5.2 *The uniformity of REAL and VIRTUAL*

A starting point to clarify these intricate results can be found in the P7 open-ended question. They seem to put forward that participants had remarkably varying feelings about the REAL condition. This disparity is also reflected in an inter-quartile range considerably larger than in the SoE scores of the two other conditions (Figure 3.5). On the one hand, many reported a weaker SoE when no virtual object was around: e.g. *"The virtual hands felt less like my hands when it was all real"*, *"The fully real environment introduced greater distance between the robotic hands and the objects"*. On the other hand, some participants appreciated REAL for its visual uniformity and often compared it to VIRTUAL: e.g. *"It seemed more 'real' when everything was virtual or real"*, or *"It felt easier when everything was either all virtual or all real"*.

Despite the relatively different distributions of REAL and VIRTUAL's embodiment scores, such comments suggest that there are similar aspects in the perception of these two conditions. One possibility is that the unmixed objects of both conditions conveyed a shared feeling of homogeneity which was partly disrupted in MIXED (cf. Section 3.5.3). This uniformity in the objects is maybe what participants referred to when comparing REAL and VIRTUAL conditions. The proportion of participants who were sensitive to it might have led VIRTUAL not to stand out as significantly different from REAL in this particular experiment. However, at this point, we can't settle whether this observation is specific to our experiment or not. A larger study will have to be conducted in order to conclude the differences between the SoE in environments with the properties of VIRTUAL and REAL.

3.5.3 *MIXED responders and rejecters*

Like in the REAL condition, the MIXED condition generated a large spectrum of responses to P7. Some expressed their sensitivity to the visual contrast between real and virtual contents: e.g. *"I was much more aware of the distinction between real and virtual objects."*. These participants often described that they saw the objects as belonging to "separate worlds" or saw "superimposed levels of reality" that disturbed them or that they disliked. Meanwhile, other participants were not bothered by the mix and even appreciated it: e.g. *"It took me a moment to notice that there were both real and virtual objects"*; *"Seeing a mixture of real and virtual objects helps to merge the two. It is harder to do the merge without it"*.

In the face of these diametrically opposed comments, it seems that some participants could be categorized as "responders" or "rejecters" of the MIXED condition. Rejecters would correspond to users who were bothered by the mix of objects and who resented the contrast between the real and virtual contents. Maybe these participants related with the real content undividedly, hence preventing them from strongly connecting with the avatar. The VIRTUAL condition could have been experienced as less disruptive than the MIXED and REAL conditions for them. For responders, on the other hand, the virtual content of VIRTUAL maybe appeared as belonging to a separate layer, overlaid on the real environment but not mixing with it. Similarly, the virtual hands of the REAL condition could also have stood out as extraneous or out of place, being the only virtual content present in the scene. Mixing objects in MIXED therefore could have helped them to make the real and virtual layers "miscible".

Altogether, the perception of the objects seems to divide the participants into three groups: those who found the mix helpful for their embodiment experience, those who found it disruptive, and those who found the absence of mix helpful (cf. Section 3.5.2). These groups seem to match the groups of rankings that came out of question P8, asking participants to rate the conditions by order of prefer-

ence (cf. Section 3.3.3). We attempted to investigate the significance and intersections of these groups with cluster analysis, but the small number of participants in our experiment did not allow us to identify them reliably. A larger study will be required to verify these theories.

3.5.4 *Personality and expectations*

So far, our results show that participants had considerably varying reactions to all three conditions. The dominant reaction led MIXED to produce stronger body ownership than REAL, but the reasons for this are difficult to define precisely. Among them, we previously identified that the perception of the virtual-real mix had a key role in the strength of the illusion. One possibility is that this perception has something to do with subjective expectations of the technology. Indeed, participants who reported a preference for VIRTUAL or REAL often commented that they experienced technical inconsistencies more strongly in MIXED (e.g. *"The textures and the brightness seemed unnatural"*). Meanwhile, the ones who preferred the MIXED condition shared their appreciation of the lighting simulation and tracking at the same level as the other conditions (e.g. *"The shadows of virtual objects made the experience quite realistic"*).

Surprisingly, however, most of them (indistinctly) expressed that the mix of objects in MIXED did not influence the perception of their avatar. This suggests that they were not conscious that their answers were influenced by the mix of content they observed.

As previous research showed correlations between personality traits and the SoE (Dewez et al., 2019; Jeunet et al., 2018), we suspected that such inner judgment could be related to the LoC profiles described in Section 3.3.1. The LoC was also formerly identified as linked to control beliefs when dealing with technology and therefore could have influenced the agency and acceptance of the avatar (Beier, 1999). We investigated whether the profiles of our participants could be predictors of the "rejecter" and "responder" categories by applying a Spearman correlation test between the LoC dimensions and the embodiment scores. While we did find several significant correlations between locus scores and embodiment scores, we could not collect enough evidence supporting such classification or allowing us to conclude the influence of the LoC.

3.5.5 *"Wearing" or "being" the avatar*

In our experiment, high appearance scores were translated to a strong personal identification with the avatar's visual traits. This visual identification is usually not explicitly cited as one of the principal sub-components of the SoE (Kilteni, Groten, et al., 2012). However, it is most often recognized as an important top-down influence factor of the feeling of body ownership (Lugrin et al., 2015; Waltemate, Gall, et al., 2018). Several studies in fact found that BOIs tended

to be weaker when the embodied object was morphologically too different or in a too inconsistent pose (Costantini et al., 2007; Lin et al., 2016; Tsakiris and Haggard, 2005). In our case, the embodied hands had a human morphology and were aligned with the participant's hands. Yet, from all scores, appearance was the one with the lowest outcomes in all conditions. Its median value hovered around 0 and spread out between -1 and 1 for about half of the population, hinting at some hesitation.

This is very likely related to the texture of the virtual hands that gave them a robotic appearance. We chose a non-human avatar to avoid bias linked to user gender, ethnicity, or general criticism caused by self-comparison (Gilbers, 2017; Lira et al., 2017; Schwind et al., 2017). However, this choice led several participants to verbally report being estranged from the virtual hands and having the sensation of "wearing" them as if they were "gloves". Indeed, personalized textures that appear natural-looking or matching the user's skins were previously shown to induce a stronger embodiment than generic textures (Haans et al., 2008; Lira et al., 2017). The lack of resemblance with the robotic skin therefore probably limited the extent of the self-identification with the avatar hands. The conformity in pose and morphology, on the other hand, seems to have positively moderated the appearance scores. This would be in line with the work of Lin et al. (2016) showing that robot hands produced a weaker visual identification than with realistic human or zombie hands, but that they could still evoke an SoE.

On another note, appearance was found to correlate to several aspects of the environment's perception: participants who were able to relate to the avatar's appearance tended to be more immersed (P1), to accept the virtual content more easily (P2, P4), but were also more bothered by the presence of real content (P5). This last correlation occurred with the ownership scores as well. We believe that these correlations denote limitations regarding the realism of the virtual content produced by our experimental setup. Indeed, as mentioned in Section 3.5.3, several participants sometimes described that their experience of the avatar was affected by the objects: the mix of virtual and real items raised identifiable differences in the perspective or lighting they displayed. The presence of real objects probably emphasized that the avatar was not real, being intuitively associated with the same nature as the virtual objects. Perhaps the enhanced awareness of this virtuality was all the more disturbing when the appropriation and identification with the hands were strong, being somehow contradictory to these feelings.

These results further highlight how subjective embodiment experiences can be. Few works have looked into how to compensate for the effects of this subjectivity. Faced with this observation, we decided to investigate strategies to mentally prepare the user to accept their avatar before the embodiment. We present our efforts to come up with such strategies in Chapter 6.

3.5.6 *On mental workload*

We found a moderate negative correlation between the total SoE score and the workload required by the VM task. However, no significant difference appeared between the three conditions in terms of mental effort, evaluated by P6. This would be in line with previous research led by (Škola et al., 2016) who compared the cognitive workload induced by RHI in VR, AR, and real settings with EEG and NASA Task Load Index questionnaire. This is encouraging for future work intending to explore virtual embodiment scenarios in real environments as the mix of real and virtual objects is often a desired feature of AR applications.

3.5.7 *Analysis discrepancies*

As described in Section 3.4, we could not strengthen the impact of our paper by reanalyzing our data with the updated version of the embodiment questionnaire or with the analysis of the first trials (Peck and Gonzalez-Franco, 2021). The outcomes of our separate analyses can be summarized as follows: (i) the 2018 questionnaire score evaluation shows there is a significant difference between REAL and MIXED in terms of Ownership sensations, (ii) the per-item analysis of the 2018 and 2021 questionnaires confirm this result from a significant difference found in Q1, and (iii) the 2021 questionnaire score evaluation and first trials analysis revealed no significant differences.

The lack of significant results in (iii) can perhaps be explained by the score computation method of the updated questionnaire and by the small number of participants. Indeed, the 2021 version of the questionnaire differs considerably from its 2018 version:

- 9 questions were removed from the original 2018 questionnaire, 4 of which were used in our initial analysis (Q2, Q3, Q5, Q6).
- Among these 4 questions, 2 were assessing body ownership in our questionnaire (Q2, Q3) initially using 3 items to evaluate this dimension.
- Instead of each question contributing to a single dimension, all questions contribute to several dimensions, making their scores highly interdependent. Ownership is now computed with items Q1, Q4, Q7, Q10, and Q11.
- 4 dimensions are evaluated instead of 6: Ownership, Response, Appearance, Multi-Sensory. Tactile Sensations (not calculated here) and Location dimensions were merged into the "Multi-Sensory" dimension, and Agency was integrated into the "Response" dimension.
- The Total Embodiment score is no longer calculated with a weighted coefficients formula, but with a simple average of all scores.

Although Q1 was not removed from the Ownership scoring of the 2021 version, its combination with items Q4, Q7, Q10, and Q11 seems to have buried the information that it was yielding (and not when combined with control statements Q2 and Q3).

In their paper, the authors of the original questionnaire found that the new version is more sensitive to SoE variations but gave similar embodiment results than the previous one (Peck and Gonzalez-Franco, 2021). This led them to conclude positively on the forward and backward compatibility of their versions. However, unlike them, we found that the two versions could give different results.

There appear to be several explanations for this discrepancy. We mainly suspect the increased sensitivity of the updated questionnaire combined with our small number of subjects to be the reason for this change. Peck and Gonzalez-Franco report that their revision amplified the dynamic range of the scale by reducing the number of items and embodiment dimensions. This reduction successfully maximized the inter-individual variability, but it may also have smoothed out the intra-individual variability across the dimensions.

From the revised questionnaire's paper, it is not clear whether the sensitivity of the scale to intra-variability has become more critical. Regardless, reporting the differences we found seems of great importance for the community as they call for further discussion on the compatibility of the questionnaires. Future research will also need to be cautious when examining the results of previous work using the 2018 and 2021 questionnaires as they may produce different conclusions.

3.5.8 Limitations and future work

Our study provided a consistent and replicable way to study environmental factors of BOIs and has raised the possibility that the world content may impact embodiment experiences in OST. However, the extent of this impact could not be addressed in its entire complexity and our analysis has to be read in the light of several limitations. First, as raised by several participants, the realism of the virtual content seems to have been limited by our rendering system. We attempted in providing physically correct parallax and reproduction of the lighting with custom implementations, but inaccuracies may have impacted the general embodiment experience. To avoid bias, we recommend future studies pay special attention to the lighting coherence of the virtual rendering. Physics simulation of real-world settings is still at an experimental stage in most AR frameworks, but taking advantage of them could be an efficient solution for this.

Secondly, the aspects measured in our post-experiment questionnaire (e.g. immersion, workload) lacked control items for most. We decided to limit the number of items to reduce the length of our experiment, but we acknowledge that including separate questionnaires to evaluate these dimensions individually would have increased the robustness of our analysis. Similarly, the secondary anal-

ysis we made following the new guidelines of Peck and Gonzalez-Franco (2021) shows that our results depend on evaluation methods and we could not confirm their validity. We believe that it is important for the community to be aware of what such changes put at stake and to investigate their impacts in future work.

Regarding our experiment's design, we would like to draw the reader's attention to the scope of this study. The purpose of our experiment was not to make a comparison of the SoE in immersive and non-immersive settings (e.g. VR versus AR) but to examine its variations within the specific context of OST experiences. This choice led us to design conditions where participants were in the real world at all times and could partly see their real hands when interacting. This situation is usual in most widespread OST headsets such as the Microsoft Hololens or Magic Leap as they render virtual content with transparency and latency.

Studies comparing BOIs in AR and VR were previously conducted by Škola et al. (2016) and Wolf, Döllinger, et al. (2020), but the impact of aspects specific to OST displays had yet to be evaluated when we ran this study. We looked at the impact of displaying different amounts of real and virtual objects together, which is a common situation in AR experiences. In the continuation of this study, it would be interesting to analyze how the SoE evolves during direct manipulation of real and virtual objects with virtual hands and to investigate the influence of the real body's visibility. Reproducing this experiment with different kinds of objects (e.g. realistic, animated, tangible) could also be an interesting avenue to obtain more complete insight into the influence of their presence.

3.6 Conclusion

This chapter presented an experiment exploring the effects of mixing virtual and real content on the SoE in AR. We evaluated embodiment illusions within three AR environments visualized with an OST setup: one displaying only virtual objects, one displaying only real objects, and one mixing both kinds. We found that users tended to feel stronger ownership of virtual hands when viewed in the presence of both virtual and real objects mixed, as opposed to when the virtual hands were the only virtual content visible. Additionally, we identified potential correlations between the ownership of the avatar's hands and user immersion as well as the perception of the virtual content's 3D registration.

Our results suggest that the content of mixed environments should be taken into account during embodiment experiences. However, we could not conclude with confidence on the origins of the observed differences and their extent to larger populations. This experiment should nevertheless encourage the community to further investigate the idea that the avatar itself is not the only moderator of BOIs. Extended research is therefore needed to clarify the environment's influence on the SoE, but also to exploit it for stronger BOIs.

If the doors of perception were cleansed everything would appear to man as it is, infinite.

William Blake

4

Exploring third-person embodiment to extend real-world perception

This chapter explores how to control avatars in the 3PP through AR. We propose the concept of using such avatars to extend our perception of physical areas ahead of ourselves. We discuss concrete application scenarios and propose interaction techniques to manipulate avatars in the 3PP as a part of a larger conceptual framework.

Contents

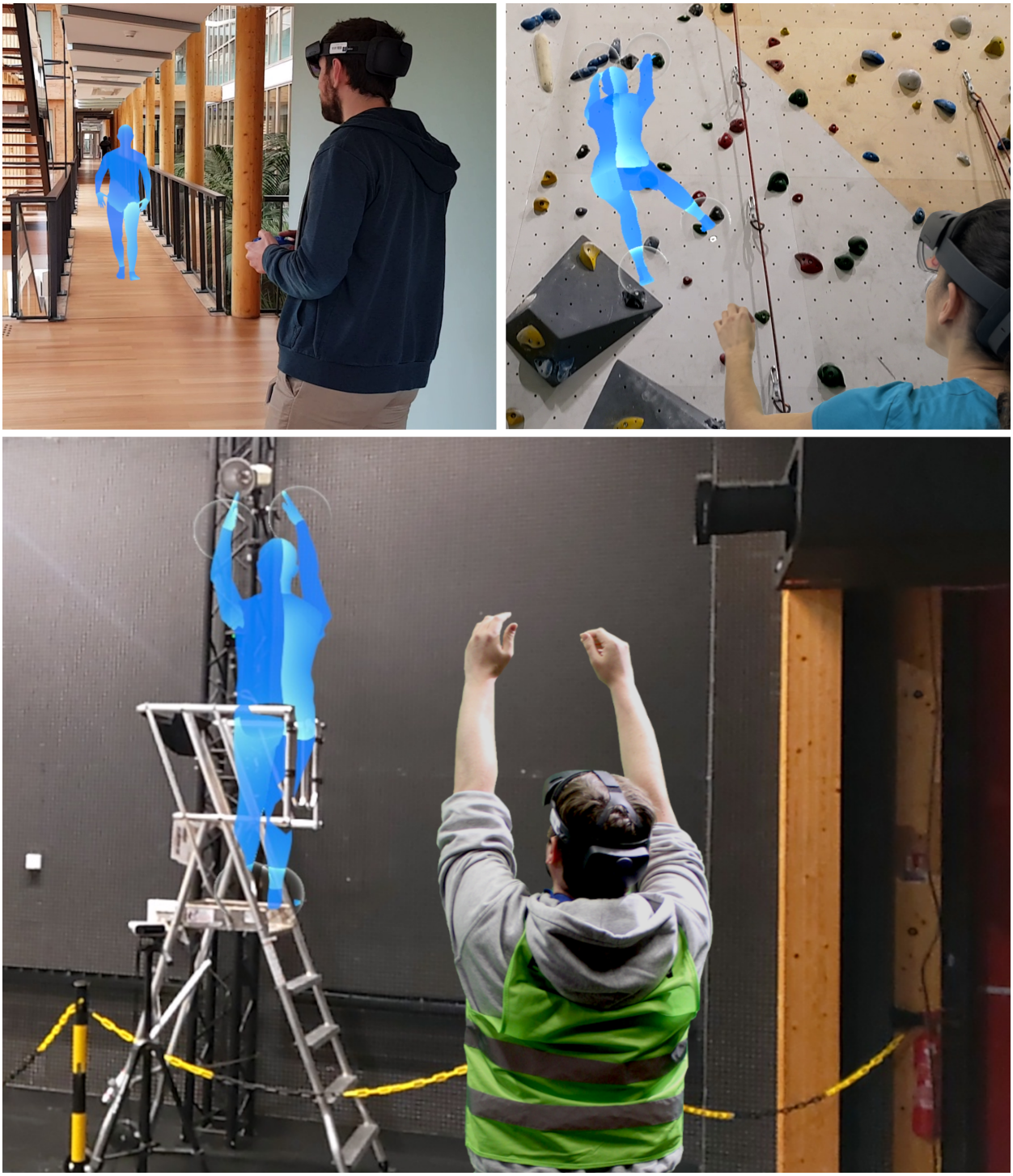


Figure 4.1: Control of self-avatars visualized through an AR headset to better perceive interactions and affordances in the physical surroundings. *Left* A user is testing fire exit paths with a gamepad. *Center* A user is planning a climbing route by gesturing. *Right* A user is evaluating possible actions on a step stool with body-tracking.

Humans perceive the physical world through action (Gibson, 2014). By moving their bodies, they provide their sensory organs with continuous access to new data that, combined with experience, allows carrying out decisions successfully. However, acting may be impossible when the environment is inaccessible, distant, or dangerous, and experience may not be sufficient to fill in the missing information. In such situations, one could wish to have the ability to be free from one's bodily envelope and explore the world from a distance.

Recent progress in MR technologies offers to do so. Whereas our physical body is intrinsically limited by its material characteristics, Merleau-Ponty (1945) argues that our perception and experience of the world cannot be reduced to material properties and may therefore be extended [1]. By modifying the inputs of our perception, research has shown that MR has the potential to enable such an extension. For example, MR was previously used to extend the reach of one's arms by virtually modifying its length, or to duplicate one's body and interact with it (Feuchtner et al., 2017; Lindlbauer et al., 2018).

[1] Merleau-Ponty distinguishes the body as the "Korper" and the "Leib". The Korper is the body as a thing, reducible to its physical characteristics. The Leib is the living and perceptive body that allows acting in space.

"Extending" the physical limits of our body through MR has many promising use cases. In particular, using AR to distance one's digital body from oneself provides novel opportunities to perceive the world that surrounds us. Users could send their virtual version nearby a distant object to get a better idea of its size, or virtually execute and visualize actions performed in relation to a physical environment. Such ability also opens the door to new types of explorations aimed at better understanding the relationship between our body, motor actions, environment, and thus cognition.

However, MR research on how virtually pushing the limits of our body can enhance our perception of the real world is still preliminary. Additionally, the control of an avatar in the third person within a physical environment is neither innate nor easy to implement. Using this ability in concrete applications or fundamental studies first requires setting up the appropriate technology and providing suitable means to act through self-avatars in the real world. In this chapter, we focus on this challenge.

We start by discussing the concept of using a self-avatar in the third person to improve the real world's perception by leveraging existing cognitive mechanisms in AR. Then, we describe a concrete implementation of this concept through a system allowing one to manipulate one's avatar from a remote place through three interaction techniques: Physical Control, Puppeteering, and Body Tracking. The results build towards a new way of using MR displays to better perceive what is already present in the physical world through a "virtual twin", instead of augmenting the world with virtual objects while staying constrained to our bodily envelope. Finally, through this approach, we aspire in helping people to make better-informed decisions during real-life scenarios.

Use cases

To exemplify the possible advantages of using a self-avatar in assisting the perception of real-world situations, we present concrete scenarios. In each case, users may improve their understanding of the physical environment, test, and refine their strategies at a distance before acting for real.

Climbing. The sport of climbing requires anticipating a route from the ground by imagining one's body in a place where it is not (Sanchez et al., 2012). Identifying which holds can be grasped from a distant position can be difficult for climbers that lack experience. These climbers could use an AR system to send their self-avatar onto the wall and plan their ascent from a vantage point on the ground (see Figure 4.1, top-right). While controlling their avatar, they may anticipate which holds can be reached next by extending the virtual limbs of their avatar and trying different postures. Having the same body size as their users, self-avatars might also help to correct false affordances that occur when observing others successfully reaching holds that are too far for their own arms.

Rescue. Misperceptions of one's abilities also occur in situations that present risks or engage certain mental states. For example, people may underestimate their ability to reach objects through small apertures when feeling anxious (Creem-Regehr et al., 2013). Providing the means to test one's capabilities virtually could help correct the effects of emotions in training situations, e.g. before entering a building threatening to collapse. A firefighter trainee that is not confident about their ability to crouch under a beam could check whether their body would fit or not by sending their self-avatar in their place first. This might help them to combat misperceptions linked to their fear and gain confidence in future interventions.

Oral presentation. When preparing for a presentation, it is often recommended to rehearse in the venue where one will present to get familiar with the spatial layout. Indeed, a proper room setup is important to ensure good communication and can help relieve stress. To better visualize space, the presenter could sit in the jury's place ahead of time and observe their avatar standing where they are planning to speak. In this way, they could see from the point of view of important attendees whether their body is hiding their slides or whether the lectern is too tall to see their gestures.

Observing a piece of art. Humans often unconsciously use their bodies as a scale to estimate the properties of objects, but this fails when they are too far or seen from a particular angle. For example, realizing how tall an actor is on a theater scene while sitting far away in the hall can be hard. Estimating the size of a very large statue while standing at its very feet can be difficult too. In these situations, controlling a virtual double of oneself could provide a familiar and reliable scale to grasp dimensions more accurately. Users could bring their virtual avatar to a point of interest and observe its size in relation to that of the object from different angles.

These simple scenarios can be generalized to other real-life tasks requiring one to imagine being in a specific posture at a distant place. We suggest that combining embodiment in the 3PP and MR has the potential to relieve a part of the mental effort demanded by this process. Before presenting the system we came up with to provide such aid, the next section briefly reviews the work that enabled its conception.

4.1 Related Work

The concept presented in this chapter is based on the vision that avatar embodiment can enhance real-world perception. We first discuss the roots of this vision in cognitive science. Next, we describe how others used MR before us to improve the physical environment's perception. Lastly, we outline the research on 3PP embodiment that inspired us when designing the presented system.

Theoretical foundations: Embodied Cognition

According to the ecological theory of perception of Gibson (2014), the perception of environments is directly related to the actions that one is capable of performing within it. The term *affordance* refers to the compatibility of environmental (as perceived by the senses) and individual characteristics (e.g. size of the body). For instance, a tree branch set sufficiently high may afford walking under, but not sitting onto or stepping over (Creem-Regehr et al., 2013).

People can usually determine if an environment allows them to perform an action without having to actually try it (Creem-Regehr et al., 2013; Mark, 1987). For example, Warren Jr et al. (1987) showed that participants estimated correctly that apertures needed to be at least 1.16 times their shoulder width to be able to pass through them without having to rotate one's shoulder while standing away. Affordances can also be recalibrated to meet new skills or situations. Ishak et al. (2008) notably found that participants adjusted their decisions about whether or not their hands could fit through an aperture after having enlarged their hands. However, the study of Mark et al. (1990) revealed that such recalibration can only occur when able to move one's point of view: the capacity to adjust information and judge affordances was considerably diminished when visual input was limited to vision through a peephole or when mobility was restricted by having participants rest their heads against a wall.

The system designed in this chapter is built on these observations and seeks to take advantage of our natural ability to understand things through action and locomotion.

Improving the perception of real environments with MR

One way to improve the perception of physical space is to allow the user to access new information by letting them adopt artificial view-

points (Dede, 2009; Klippel et al., 2021). Systems implementing such viewpoints use cameras to reconstruct the environment or objects of interest in 3D, and then immerse their users in the resulting virtual environment where any perspective can be displayed (Lindlbauer et al., 2018; Miyaki et al., 2016; Roo et al., 2017)[2].

Rather than substituting the user's sight, some research proposed using situated visualization to help imagine the effect of one's actions. For example, Leigh et al. (2015) developed a mobile see-through AR system letting users see how cubes would fall if they touched them when approaching their hands during a block-building task. Other papers looked into making already existing information easier to perceive by augmenting various sensory channels, including vision, audition, and touch (Daiber et al., 2013; Everingham et al., 1998; Zöllner et al., 2011; Zoran et al., 2013). Schoop et al. (2018) notably proposed to sonify objects outside of the field of view of users to inform of their presence, type, and distance.

We draw upon these examples. Compared to previous work, we propose to let users explore and sense their real environment at will through self-initiated action in OST AR.

[2] The experience of such systems is close to VR — even in the case of Remixed Reality, a system developed by Lindlbauer et al. (2018) where real-time photogrammetry of the physical world is displayed.

Increasing spatial awareness through third-person avatars

Often used in video games, the 3PP provides a wide field of view enabling users to quickly perceive elements around them. Previous work exploiting this view in real environments usually implemented it by moving the user's viewpoint outside of their body, as in the work of Komiyama et al. (2017)[3]. Similar to out-of-body experiences, this point of view was shown to improve the spatial awareness of physical environments. In MR, Salamin et al. (2010) notably showed that moving the user's viewpoint behind their bodies reduced the training required for a ball-catching task. Liu et al. (2017) also showed that the 3PP resulted in slightly less precision during a measurement task, but it allowed for being three times faster. In VR, the effects of the 3PP are contrasting. While several papers showed positive effects on spatial awareness in various kinds of tasks, Medeiros et al. (2018) found that the advantages of this perspective vary with the type of task and the avatar's appearance. It is not clear at this point whether this also applies outside of virtual worlds.

Putting in place the equipment required for out-of-body-like experiences of the real world is sometimes impossible or inconvenient due to access restrictions, cost, or physical constraints (e.g. the environment is dangerous). In such a case, an alternative approach to providing a 3PP is to display a duplicate of the user's body ahead, observed from a first-person viewpoint. This visualization is similar to autoscopic experiences[4], as two bodies are visible. Although papers implementing such a perspective exist, they do not exploit it to increase the spatial awareness of physical spaces (Nimcharoen et al., 2018; Rosa, J.-P. v. Bommel, et al., 2019). We propose to start explor-

[3] In this case, there are no avatars: it is their real bodies that the users see in the third-person. Obtaining such perspective can be done by placing a camera (ideally mobile) at the desired point of observation and streaming its view to the user's HMD, or creating a 3D live-updated reconstruction of that space and immersing the user within it, as in (Lindlbauer et al., 2018).

[4] Autoscopic hallucinations are the highly realistic experience of seeing one's body or body parts from an internal point of view but at a distance, as in a mirror (Arenz, 2001).

ing this approach in the current chapter.

In summary, what fundamentally differentiates our work is: (i) our users do not change their natural visualization perspective, (ii) they see their real environment rather than a remote/virtual one, and (iii) they have control over the exploration of their surroundings through a rigged avatar, matching their body dimensions. To our knowledge, we are the first to explore such new directions.

4.2 *Rationale and design considerations*

Our final goal is to improve the perception of environmental properties in distant spaces to make better-informed decisions and prepare for actions. We do not seek automatic methods that could analyze the physical environment on-the-fly and optimize ideal body movements. Instead, our approach is to leverage existing cognitive mechanisms by providing people with the ability to simulate their actions outside of their peripersonal space. To do so, we propose to rely on a virtual avatar that represents the user, that is embedded in the real world, and that can be easily manipulated.

The 3D registration of the avatar in the real world requires the use of an MR system. HMDs appear well suited for our objective as users can observe virtual objects while keeping their hands free for interaction. To safeguard natural perception of the real world, we opted for an OST HMD. As explained in Chapter 2, OST HMDs provide an unmediated view of the real world that ensures the visual and proprioception information are synchronized. Current OST HMDs also have scanning capabilities that favor the consistent integration of the virtual avatar within the real environment.

Of course, for the user to perceive the real world as if they were actually experiencing it, the sizes of the avatar's limbs have to be similar to the user's body. Beyond limb sizes, reproducing the user's traits with fidelity and realism does not seem essential for this system. Realistic avatar appearances may additionally provoke Uncanny Valley effects that can negatively impact user experience (I. Wang et al., 2019). Therefore, we decided to personalize the avatar's body, but cover it with an abstract and generic texture.

Lastly, as the actions of the users in the real world should be as varied as possible, we explored and identified three main potential interaction needs:

- **Travelling and wayfinding:** First, it may be interesting to stay in place and explore possible paths in the real world by moving one's avatar from one location to another, e.g. to better perceive the dimensions of a room.
- **Posture editing:** Second, beyond global movement, individual limb manipulation may be valuable. An example is when trying to figure out which holds can be grasped before climbing onto a boulder. In this case, independent and fine control over each body limb is necessary.

- **One-to-one mapping:** Third, it may be interesting to project, through the avatar, a particular body gesture in the real world. For instance, a dancer could wish to check if they have enough space to perform a particular figure on a stage with fragile decor by virtually performing the figure at a distance, in a safer zone.

To accommodate for the variety of tasks related to these different needs, we have explored three interaction categories that are described in Table 4.1. Depending on the environment and goal, one may choose the best-suited approach, or combine them for comprehensive exploration. The choice of the interaction method may also come from users' specific needs. For example, an elderly user may have difficulties with precise motor input but may be able to control the avatar with a controller instead.

Table 4.1: Overview of the interaction modes that we have explored to allow manipulation of an avatar from a distance, in AR.

Category	Scale	Implementation	Perks	Limits
Travelling, wayfinding	World	Keypad, game controller, joystick, tactile displays, etc.	Requires minimum physical effort. Allows making the avatar walk over distances without actually moving. Can be used eyes-off after little training.	Remembering the mapping between buttons becomes difficult after only 2-3 buttons are used. Control over the avatar is limited to a set of pre-recorded animations.
Posture editing	Limb	Pinching, dragging, pointing, puppeteering, raycasting, etc.	Provides the finest control of the avatar's posture. Metaphore-based interactions are easy to learn.	It's difficult to manipulate several limbs at the same time. Gesture recognition is not always reliable and can be physically tiring. Implementing postures can be slow.
One-to-one mapping	Body	Optical, inertial, mechanical, magnetic tracking, etc.	Most direct and natural control (one-to-one mapping). Provides vestibular cues (inertia and balance) and a strong sense of agency.	Multiple technological constraints, including sensor range, cost, and portability. Noise is introduced in movements due to tracking errors. Achieving certain postures can be impossible from a distance.

4.3 Implementation

As a proof of concept, we implemented a prototype enabling users to control a self-avatar as described in previous sections. We implemented three modes to control this avatar whose limbs are matched to the size of its user. As we meant to propose several control techniques to animate the avatar, we decided to opt for a rigged mesh model rather than a point cloud avatar which only affords body tracking. Figure 4.2 illustrates the global system and its different components in use.

Julie Valcaneras contributed to the development of the Physical Control mode as a part of her internship for her engineering degree. The system's code is available at: `https://gitlab.inria.fr/agenay/ISMAR22-whatCanIDoThere`.

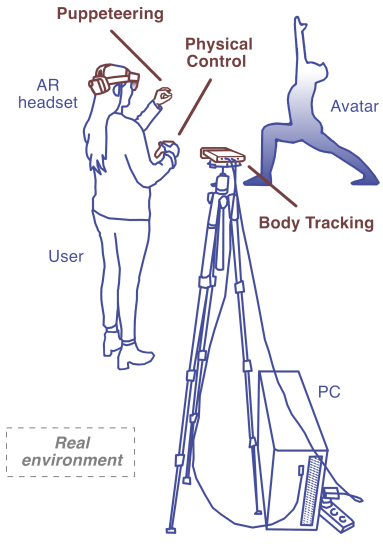


Figure 4.2: Global structure of our system. *Display* A Microsoft Hololens 2. *AR Module* Unity application running on a PC streaming it to the display over WiFi. *Avatar* Life-sized 3D mesh matching the user's body shape, registered in 3D space.

4.3.1 Display

We used a Microsoft Hololens 2 to display the avatar in OST AR. This HMD has an approximate field of view of 54 degrees diagonally and is equipped with 6 cameras, 1-MP time-of-flight depth sensor, and inertial measurement units allowing real-time surface detection, hand tracking, and positional tracking with six degrees of freedom.

4.3.2 *AR Module*

We exploit the Hololens 2 sensors in a C# implementation to register the avatar in 3D space and detect user gestures. Environment detection is also used to implement occlusion and collisions of the avatar with real surfaces. To do so, we use Unity3D 2019.4.16f1 and Mixed Reality Tool Kit (MRTK) v2.6.1. to build an application for Hololens 2. This application is also in charge of processing the user inputs of all three modes and managing the user interface.

4.3.3 *Avatar Generation*

Medeiros et al. (2018) found that mesh models resulted in lower accuracy during navigation tasks in 3PP VR compared to point cloud avatars. It is unclear whether this also occurs when exploring physical environments too. However, to avoid potential discrepancies that might have caused such negative effects, we personalize the avatar to match the user's morphology, gender, and limb sizes with the free avatar creation tool Virtual Caliper (Pujades et al., 2019) based on the SMPL model (Loper et al., 2015).

We use all 6 input parameters of Virtual Caliper to generate user-matching avatars before testing: height, weight, arm span, inseam height, inseam width, and wrist-to-shoulder distance. The generated model is rigged and skinned but does not include a body texture (hair, clothes, etc.). We used a generic abstract texture to cover the avatar (see example in Figure 4.3, 4).

4.3.4 *Control modes*

We implemented three control modes, each fitting in one of the usage categories described in Table 4.1: the Physical Control mode is for traveling and wayfinding, the Puppeteering mode is for posture editing, and the Body Tracking mode is for one-to-one mapping. Figure 4.3 illustrates the usage of our system, from the selection of a mode (Puppeteering) to its application in a climbing gym.

Physical control mode

For this mode, we used a wireless XBOX controller (X/S series). It was paired in Bluetooth to the Hololens 2 and its button mapping was managed by a Unity application. Since this mode is dedicated to providing navigation, buttons were mostly mapped with actions linked to locomotion through pre-recorded animations. The button layout we chose follows conventional controls of western platform games: left joystick for moving and turning, (A) button for jumping upwards, left trigger button for crouching, and (X) for leaping. We additionally use (Y) for extending arms in T-pose and the down pad button for laying down. When not moving, the avatar was animated with an idle animation making it appear to breathe slowly.

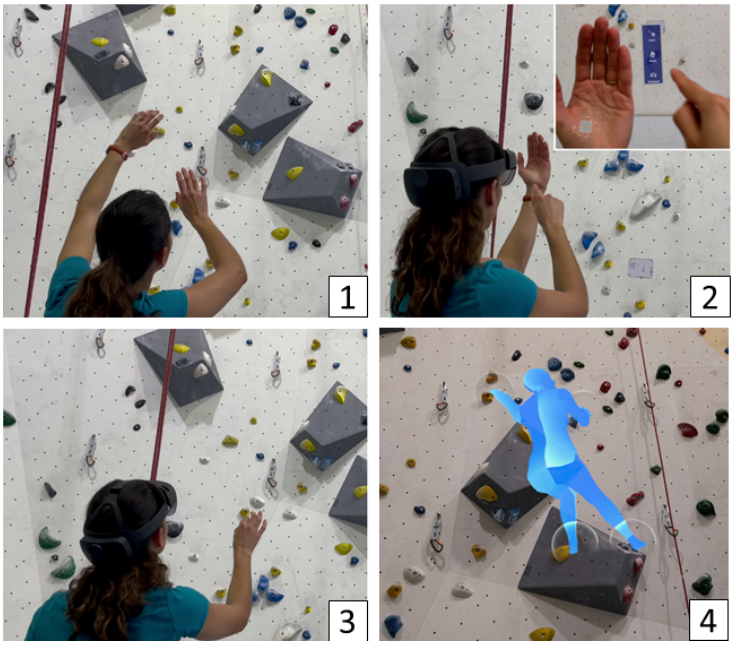


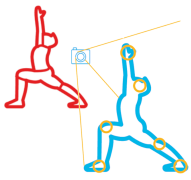
Figure 4.3: Example usage of our system. 1) A user is trying to figure out how to reach a hold on a boulder. 2) To better visualize her possibilities, she puts on an AR headset and launches the Puppeteering mode. 3) She then sets the position of her avatar with hand interactions. 4) The avatar as seen by the user (photo shot from the headset).

Puppeteering mode

In this mode, the avatar's behavior is set to that of an idle active ragdoll whose limbs can be moved by dragging around transparent spheres attached to them (see Figure 4.3). These spheres respond to selection and manipulation through input gestures (pinching, dragging, and raycasting) detected by the headset. To implement this, we use MRTK and IK scripts. Colliders and joint limits of the avatar's bones are generated automatically with the help of the PuppetMaster package (v1.1) (Pärtel Lang, 2022). We let the user drag the avatar's position in space without affecting its posture by selecting its body. They may also rotate it or its individual limbs by making a twisting movement with their wrist. To facilitate the placement of the avatar in space, we froze the body's rotation to only the vertical axis by default. To enable other rotation axes, users may press a "free rotation" button attached to their right hand that appears exclusively in this mode. Lastly, this mode also includes a "Reset" button (also on the right hand) that puts the avatar up-straight in T-pose, 1 meter ahead of the user.

Body tracking mode

The body tracking mode employs a Microsoft Azure Kinect to track the position and rotation of 32 body joints. Tracking data is extracted with Microsoft's Body Tracking SDK (v.1.1.0), and streamed to the Hololens 2 through a PC (NVIDIA GeForce RTX 2080, Intel Core i9-9900K, 32 GB of RAM). To avoid having to build a client-server network, we used the Holographic Remoting tool provided by MRTK. There was a total of 32 joints streamed by the tracker. For the tracking to function properly, the user needs to stay in the camera's sensor range. The user may face any direction while in this mode,



but body parts that are not in the field of view of the Kinect cannot be tracked. For this reason, the camera should be placed in a manner that minimizes self-occlusion.

The resulting architecture is illustrated in Figure 4.4. Depending on the task, the best-suited mode can be chosen. One can also use a combination of the three modes by switching between the control modes through a virtual menu attached to one's hand. The setup is flexible, so depending on the use case, one, two, or three of these modes can be operated.

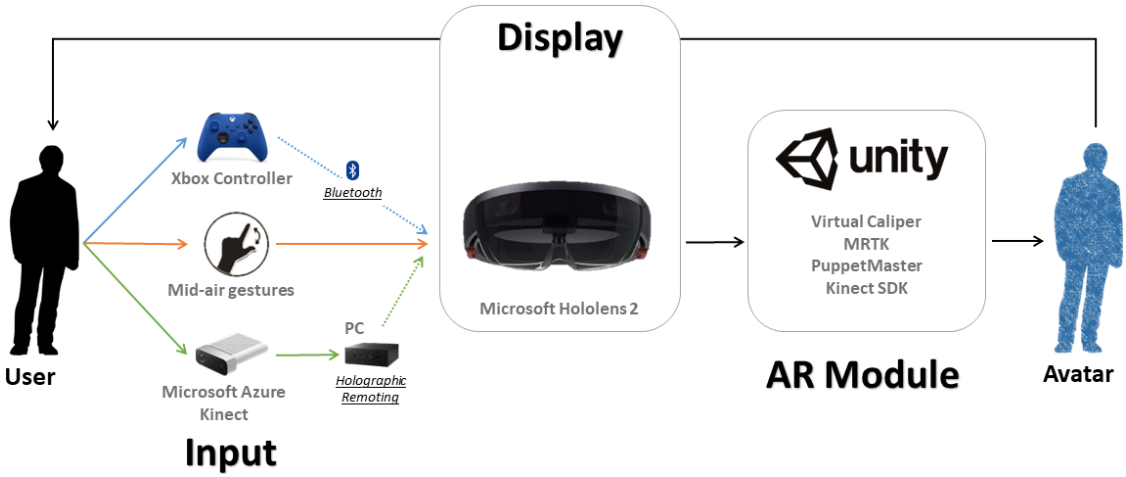


Figure 4.4: Global architecture of our proof-of-concept system.

4.4 Discussion

We investigated how to take advantage of 3PP virtual embodiment to access locations without physically entering them. To explore this idea, we implemented a proof-of-concept giving full control over an abstract duplicate of oneself through three techniques. The controlled avatar constitutes a visual reference that can be used as a means of comparison and simulation to better understand one's environment. Designed to adapt to various situations, the control modes we propose are three examples among others. While presenting them, our goal was to illustrate and awaken the interest in using avatars to better perceive the real world through MR.

Indeed, we believe that their usage can benefit users in a considerable range of scenarios. For example, perceiving the actual size of distant objects or people becomes easier, as does understanding how one's body occupies space. Practicing movements and iterating on errors may become simpler during sports training or skill learning. Users can also use this ability to evaluate action possibilities by simulating them from different standpoints. After visualizing each option in the situation, choosing the best is made easier. This can additionally be profitable during authoring: new movements or body poses may be designed (or "sketched") and edited in real-time iterations,

e.g. when designing a choreography or a play directly on the stage.

As such, the multimodal system we present has several limitations. First, it requires multiple pieces of equipment (PC, Kinect, Hololens, XBOX controller) which make the transportation and installation of the system not so easy. To minimize this trade-off, the equipment we chose is all off-the-shelf and we worked on making them work independently. This means that part of them can be discarded without problem when it is impossible to set them up for contextual reasons (e.g. installing a computer is impossible). Modes that do not use discarded equipment will still be functional and the system can still be used, although not to its full potential.

Secondly, simulating body movements and postures successfully does not guarantee that users will be able to accomplish them with their bodies. The example of climbing a boulder shows well that uncertainty remains: another climber could block the user's trajectory, or a hold could be too slippery to grasp. Additionally, it is left to the user to judge whether their stamina, muscle strength, and balance skills will allow them to achieve the climb they are simulating. Although uncertainty remains, having a concrete visual example of possibilities pertaining to one's body proportions may nevertheless help in the process of evaluating environmental properties.

4.4.1 Future work

Finally, the system we implemented can be expanded with other modes and improved with countless other techniques to adapt to various and specific situations. Creating an ultimate system implementing a myriad of modes was outside the scope of the present study, but we seek to bring light to the many possibilities that are available for future work to explore. The following list provides some examples of features that could complement the core system we proposed:

- **Record and rewind feature.** It can be difficult to observe the avatar's body movement if one has to perform the same movement at the same time. Therefore, it might be useful to record and playback this movement at a different pace as previously proposed in Remixed Reality (Lindlbauer et al., 2018).

- **Contact feedback** As suggested by some participants, identifying when the avatar is colliding with physical surfaces hidden behind the avatar itself can be hard. We imagine that the inclusion of haptics could allow the user to feel surfaces that the avatar touches remotely. Visual highlights or sound cues could also be implemented (Schoop et al., 2018; Zöllner et al., 2011).

- **Affordance detection** Areas of the environment that afford specific actions could be highlighted and labeled (e.g. "grabbable", "walkable") to help the user visualize all possibilities at once.

- **Flexibility and strength calibration** By measuring the user's flexibility and strength, one could more precisely calibrate the avatar to the user's body capabilities. This would allow the implementation of more accurate physical constraints when manipulating the avatar.
- **Duplication of the avatar** It could be interesting to allow the user to manipulate several copies of their avatar all at once, or individually e.g. in authoring scenarios where the user needs to have a side-by-side visualization of different actions.

4.5 Conclusion

In a nutshell, this chapter sought to present the concept of using 3PP avatars in AR to extend our perception of the space of possibilities in the real world. The produced implementation may serve as a starting point for future research aiming to explore such usage of avatars in AR. Before evaluating the usability of the presented system and investigating other techniques, we propose to first verify that 3PP embodiment may effectively help accomplish what we aspired to do, i.e. increasing our understanding of the environment around us and action possibilities. Chapter 5 presents a user study evaluating this idea while using the Puppeteering mode.

The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them.

Sir William Bragg

5

Evaluating the effects of third-person embodiment in AR on real-world perception

The previous chapter presented our idea of using 3PP avatars in AR to explore the real world remotely. We now describe a user study examining how such interactions may enhance our perception of affordances in physical places ahead of us. We ran two experiments using the Puppeteering mode, and one follow-up experiment using the Physical Control mode presented in Chapter 4.

Contents

In the present chapter, we use the 3PP embodiment system described in Chapter 4 to verify whether self-avatars may effectively help to improve affordance perception in physical environments. To do so, we ran three experiments:

- **Experiment 1**: we evaluated whether seeing and manipulating a 3PP self-avatar *in situ* could effectively increase the user's accuracy when judging their action capabilities in physical space, as theorized in Chapter 4. We used the Puppeteering mode presented in this chapter to provide control over the avatar.
- **Experiment 2**: within two different environments, we evaluated whether one's perception of real environments could also be improved *after* having used the system, and not only while using it. We also used the Puppeteering mode for this experiment.
- **Experiment 3**: this follow-up experiment further assesses how the experience acquired by controlling the avatar is memorized and subsequently used once the user can no longer see their virtual duplicate. We used the Physical Control mode for this experiment.

Our results validate our hypothesis that using 3PP avatars can improve affordance perception in areas that are outside of our proximal zone (Experiment 1). But once the avatar is no longer here to help, we could not show that the experience acquired through it helped during new affordance judgments (Experiments 2 and 3).

5.1 Related work

Previous research has found that people can correctly estimate the affordances related to other people and that it can serve to judge affordances related to oneself (Creem-Regehr et al., 2013; Mark, 2007). This perspective-taking ability also works with virtual avatars, as shown by multiple studies (Arend et al., 2021; Böffel et al., 2018; Salm-Hoogstraeten et al., 2021). In VR, Saxon et al. (2019) notably found that users responded faster during a Simon task[1] when a generic virtual body was present ahead of them, without being embodied by the user. Their results illustrate the robustness of the human ability to internalize the perspective of others.

[1] The participants had to judge whether a ball was reachable with their left or right hand if they were sitting on the chair ahead. The ball appeared either on the right or left side of the chair at different distances.

Rather than providing a generic non-embodied avatar, we customize its appearance to match its user's morphology and allow the participants to control it. We hypothesize that this control will even further benefit the user when assessing their action possibilities as they may adapt the avatar's posture to simulate actions and obtain direct feedback, consistent with their body's shape and size.

Instead of measuring the user's time performance in one particular task, we propose to evaluate the accuracy of their judgments over various actions. We additionally collect subjective feedback on their confidence during their estimations and the help that they felt their avatar provided.

5.2 Experiment 1

This first experiment compared the accuracy of users when estimating their ability to perform certain actions in a physical environment over two conditions: one where they simply sat and observed the place where the action was to be performed, as they would normally do, and one where they could send their self-avatar to simulate these actions remotely through AR. We hypothesized that seeing an avatar representative of oneself in a distant place would allow acquiring information that could otherwise be hard to imagine.

5.2.1 Apparatus

System. We used a Microsoft Hololens 2 to display the avatar. The experiment's system was essentially a simplified version of the AR embodiment system described in Chapter 4, but containing only the Puppeteering mode. Indeed, before comparing the effects of each control mode, we wanted to verify whether a self-avatar could effectively enhance the perception of affordances in general.

We did not choose the Physical Control mode because remembering the button mapping becomes difficult when numerous actions are required. The Body Tracking mode also seemed inadequate as it made some actions impossible or uncomfortable to do by nature (e.g. sitting in the air) or due to sensor range limits (e.g. walking, lying on the floor). On the other hand, the Puppeteering mode offered the flexibility to test various and complex postures without requiring a long user training (see Chapter 4).

Avatar. We generated life-sized rigged mesh avatars for each participant of the AR condition. To do so, we collected their gender information and body measurements (height, weight, arm length, inseam length and width, and wrist-to-should length) and applied them in Virtual Caliper before the experiment. The subjects were informed that this data was used to generate their avatars. Like in Chapter 4, we used an abstract texture to cover the resulting mesh as representing the user's identity was not necessary for our purpose.

Environment. The experimental environment we set up to test the participants' judgment abilities was designed to afford multiple types of actions. It consisted of a spatial arrangement of diverse objects (blocks, climbing holds, chains), laid onto or above a platform made with tables as illustrated in Figure 5.1. Because the participants had various body sizes, we adapted the placement of these objects so that the difficulty would not vary across users.

Figure 5.1: Experimental environment of Experiment 1, as seen by the participants. It was revealed to the participants at the last minute.

5.2.2 Methodology

To test whether manipulating avatars could lead users to be more accurate when estimating what they can do at distant locations, we used a between-subjects design with two conditions.

- **Condition "R" (Reality)**: the participants had to answer questions asked orally by the experimenter about their ability to perform actions (e.g. touching an object) or about where some body parts would arrive if they were in specific postures, at specific locations of the experimental environment ahead of them (see Figure 5.1).

- **Condition "AR"**: the participants had to simulate the same actions and postures that were inquired about in condition R by manipulating their self-avatar, displayed in AR. After each action, they had to answer the questions respectively asked in condition R while seeing their avatar in place and still being able to adapt its posture (see Figures 5.2 and 5.3).

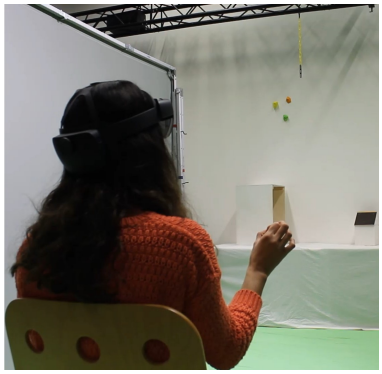


Figure 5.2: Participant manipulating their avatar during the AR condition.

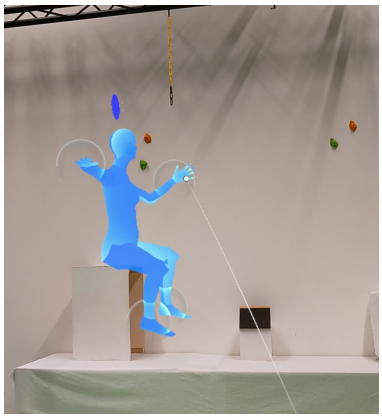


Figure 5.3: Example view of a participant while manipulating their avatar with ray-casting.

The AR condition was preceded by an eye calibration for the AR headset and by a short training session (10 minutes). During this training, the participants were presented with their self-avatars and instructed to manipulate it as dictated by the experimenter. They were also asked to observe its similarity with their body shape and size by walking around it and comparing the lengths of their limbs.

Before starting, the participants were invited to sit on a chair placed 4 meters ahead of the environment of the experiment, hidden by a large panel. The experimenter would then ask them to close their eyes and uncover the environment. In the case of the AR condition, the experimenter would also take the AR headset back while the user still closed their eyes and scan the environment again to ensure it was properly detected by the Hololens 2. They then returned the HMD to the participant (if they had any), sat out of sight, and instructed them to open their eyes again.

After going through their conditions, the participants were invited to fill out a questionnaire assessing their subjective experience of the system (see Section 5.2.3). The whole session lasted about 45 minutes. The total hands-on time with the AR system in the AR condition was of about 25 minutes during which the experimenter could see the participants' viewpoint and their interactions with the avatar via a live video stream from the HMD.

5.2.3 *Measures*

The questions asked after each action aimed to evaluate the participant's judgment of affordances and their accuracy when mentally projecting themselves at the observed locations. They were of the following types:

- **Yes-or-no questions**: "By looking at [*some place*], do you think you could [*do some action*] if you were [*located in some place, in some posture*]? What is your level of confidence from 1 to 10?". Example of answer: "— Yes, 8"

- **Estimation questions**: For these questions, we projected a scale with linear, non-standard units on the wall behind, as illustrated in Figure 5.2.3). "By looking at [*some place*], which unit do you

think [*some body part*] would reach if you were [*located in someplace, in some posture*]? With what margin of error?". Example of answer: "— Unit 31, plus or minus 2 units".

Appendix A.1.1 provides the full details on these questions. We did not compare the user's accuracy when using their avatar to when using other measurement tools (e.g. AR ruler). Indeed, our objective was not to measure distances remotely, as we could do with a telemeter, but really to better understand the extent to which an externalization of our body may help us to perceive our possibilities for action in distant physical surroundings. Similarly, we did not measure completion time as it was out of our research scope.

The subjective questionnaire we applied contained 11 items on a 7-point Likert scale and 5 comment boxes letting participants write their thoughts on the avatar's appearance, integration in the real world, control, the help it provided, and their general appreciation of the system. Full details are also available in Appendix A.1.2.

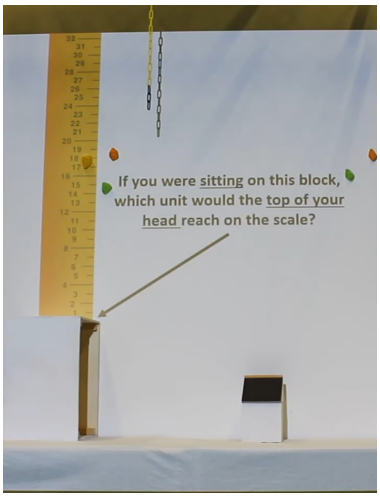


Figure 5.4: Example of estimation-type question that we asked. Participants had to answer with a unit and with a margin of error on a scale projected on the wall. They were informed that the units were arbitrary.
The projected text says: "If you were sitting on this block, which unit would the top of your head reach on the scale?"

5.2.4 *Participants*

We recruited 18 participants from age 22 to 45 ($\overline{m} = 30.5, SD = 8.8$, 10 identified as males). Among them, 10 had never used an AR headset before. No compensation or course credits were issued, and all participants were unaware of the purpose of the experiments. Before the testing session, we collected the body measurements of the volunteers to generate their avatars.

5.2.5 *Results*

To better understand the inherent strengths of this system for judging affordances and projecting oneself mentally, we assessed three main aspects: (i) the quantitative accuracy of user judgments, (ii) their level of confidence during evaluation, and (iii) the qualitative appreciation of the system. The first two were evaluated from the answers to the evaluation questions for each separate study, whereas the last was evaluated with the subjective questionnaires. One participant was removed from the analysis of each study due to incorrect body measurements that impacted the avatar's perception.

Accuracy assessment

To evaluate errors, we used a theoretical ground truth for each question. This ground truth was obtained by using the body measures that participants had provided for Virtual Caliper and complementary measures taken at the end of the experiment. We computed individual scores for each condition and question type.

- The scores of the yes-or-no questions were computed by averaging their answers, coded with 1 or 0 (for true or false). Scores ranged from 0 to 1, with 1 meaning that all answers were true.

- The score of the estimation questions was computed by averaging the difference between the participant's answer and the ground truth (i.e. the unit they should have been able to reach). A low score means the participant was accurate.

We then performed descriptive and statistical analyses of these scores. Shapiro-Wilk tests show that the scores did not follow a normal distribution, so we used non-parametric tests to evaluate the significance of differences.

Descriptive analysis. The mean scores of the yes-or-no questions were 0.73 for the R condition and 0.95 for the AR condition. This means that the average success rate was close to 100% in the AR condition. The mean scores of the estimation questions were 4.93 for the R condition and 2.01 for the AR condition. Being closer to 0, the mean of the AR condition indicates that participants made smaller errors than in the R condition. Looking at the error values, it appears that the participants of the R condition tended to underestimate their body sizes, whereas those of the AR condition were closer to the ground truth and sometimes slightly overestimated their sizes. These results are summarized in the boxplots of Figure 5.5.

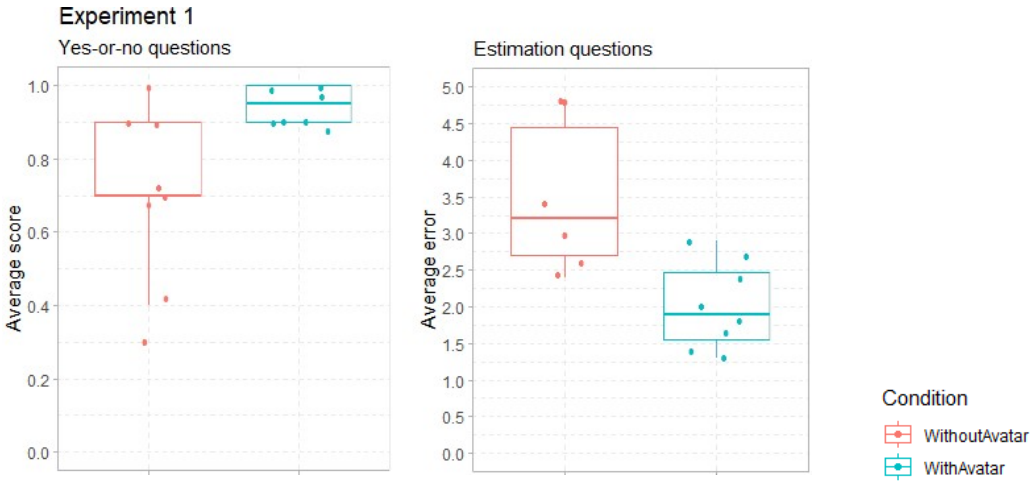


Figure 5.5: Boxplots representing scores of the yes-or-no and estimation questions asked during Experiment 1.

Statistical tests. We applied Wilcoxon rank-sum tests which showed that the scores were significantly different across the R and AR conditions for both types of questions (yes-or-no: $p = 0.049$*, estimation: $p = 0.003$**), with a **moderate** effect size for the yes-or-no questions ($r = 0.49$) and a **large** effect size for the estimation questions ($r = 0.74$). We conclude that the participants who used an avatar to answer the evaluation questions were more accurate than those who did not have an avatar to do so.

Confidence level

We ran a between-subjects analysis to compare the confidence ratings and margins of errors given by the participants in the R and AR conditions. Shapiro-Wilk tests show that the confidence data follows a normal distribution. Regardless, we still used non-parametric tests for both as the data is not continuous and the number of participants is rather small.

Descriptive analysis. The average confidence rating was 7.78 for the R condition and 8.66 for the AR condition. This means that the average confidence ratings were closer to the maximum confidence level (value of 10) in the AR condition. The average margin was 1.6 for the R condition and 1.05 for the AR condition. Being closer to 0, the mean of the AR condition indicates that participants estimated they made smaller errors than in the R condition.

Statistical tests. The results of the Wilcoxon rank sum tests show that the margin size given during the estimation questions was significantly different across the R and AR conditions ($p = 0.017$*) with a **large** effect size, but not the confidence rating of the yes-or-no questions ($p = 0.135$). We conclude that the participants who used an avatar to answer the estimation-type questions were more accurate than those who did not have an avatar.

User feedback

Figure 5.6 shows the results of the post-experiment questionnaire of Experiment 1. The answers we collected generally showed a good impression of our system.

Help provided by the avatar. We received quite positive feedback regarding the help that the avatar brought when answering the evaluation questions: 89% of the participants judged that the avatar helped them gain confidence when answering (item 7). Written comments included: "*Without the avatar, it would have been difficult to evaluate the answers to the questions*", "*The avatar is of great help*".

Trust in the avatar. Although many of the participants confessed that they were surprised by the places they could reach with their avatars, the majority seems to have trusted what they saw and relied on it to answer the questions. One participant commented the following: "*I felt I could completely trust the avatar as I had compared myself to it, so I also felt it was a good representation of me in space*", and "*felt quite sure of my answers*".

Surprisingly, two participants decided not to use their avatars to answer the evaluation questions. These two participants are those that indicated that the avatar did not correctly represent the space occupied by their own body (item 10) and that thought the avatar might be misleading (item 11). They expressed that they did not

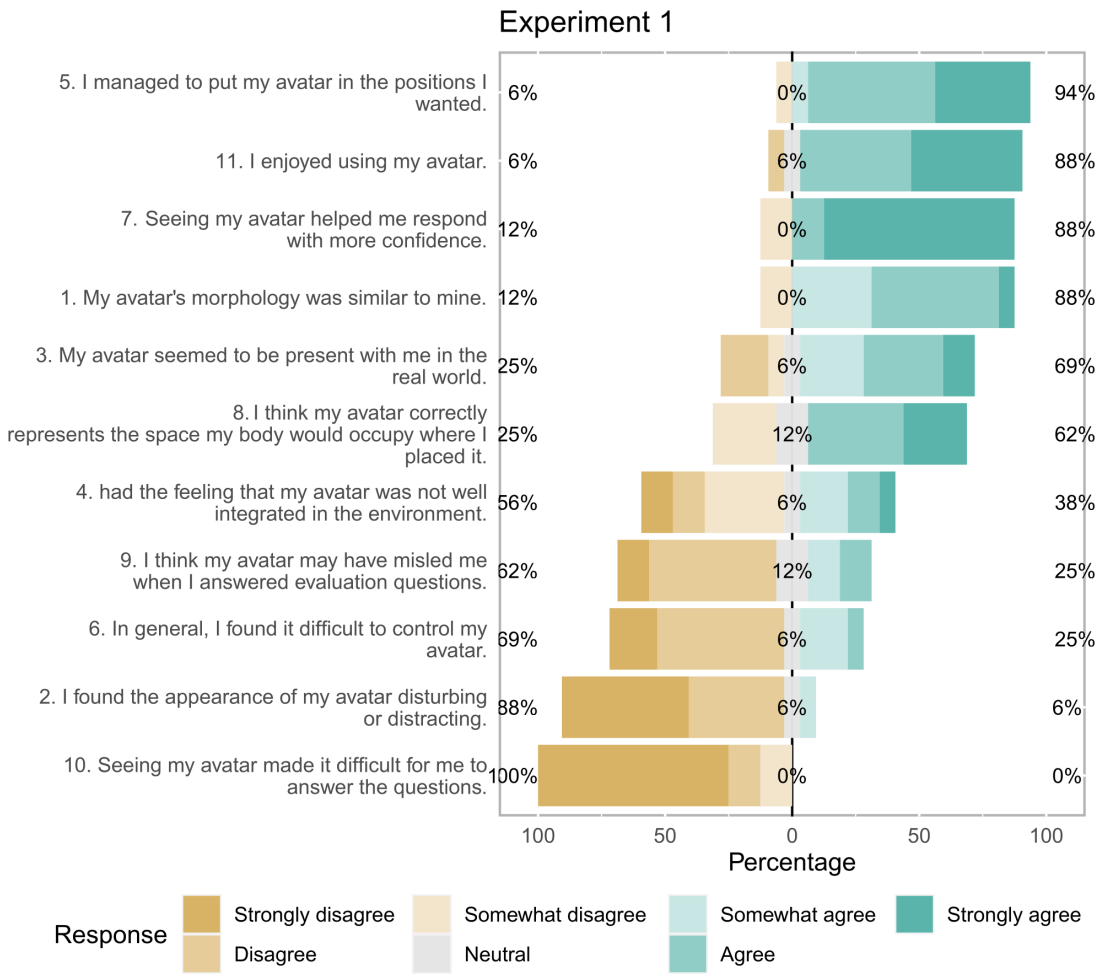


Figure 5.6: Results of the subjective post-experiment questionnaire of Experiment 1 (translated from French). Detailed percentages are available in Appendix A.2.2

trust to use it to answer the questions accurately and preferred to rely on their impression, leading them to answer with lower estimations than what the avatar indicated[2]. Still, their scores seem to have been influenced by the avatar as they were higher than those of the participants who did not have an avatar. We expect this behavior could be either due to depth perception issues that led the avatar to look closer than it was or to a general mistrust of the technology. Further investigation will be necessary to clarify this point. Either way, this does not seem to be the general tendency.

[2] During the experiment, these participants explained that their avatar seemed to have the same body size as them when they could stand up close to it during training, but that it appeared bigger when it was further away (while they were seated).

Avatar appearance. Most participants reported that the avatar's body seemed to have the same shape as their own (item 1). One participant shared the following: *" Initially, I thought she was bigger than me in terms of scale [...] but when I walked close and compared my arm length and height, etc. I felt I could confirm she was very similar to me"*. This feeling was often expressed orally during the training session by other participants. Similarly, most participants did not find their avatar's appearance to be disturbing or distracting. One participant shared that they *"enjoyed the neutrality of the appearance"*.

3D registration. The questionnaire also seems to show that the avatar was usually perceived as being well-registered in the environment. However, some participants experienced environment detection issues that led them to have more mitigated answers. Detection errors occurred when the device updated its understanding of the environment and interpreted the environment to be closer to the participant than it was. Nevertheless, 69% of the participants reported that the avatar felt "present" with them in the real world (item 3), which suggests that they were usually able to ignore the detection errors when they occurred.

Avatar control. Regarding the control of the avatar, 94% of the participants reported being successful in putting the avatar in the positions they wanted (item 5), and 69% reported they did not find controlling the avatar difficult (item 6). We collected comments such as "*quite intuitive*", "*really surprising*", "*practical*", or "*easy to take in hand*". One important cause of difficulty seems to have been more related to the HMD's limitations than to the avatar manipulation itself: participants sometimes lowered their hands too much for the headset to see them and their gestures were no longer detected. Nevertheless, one participant shared that they were *"rather surprised by the ease of use despite the limits of range sometimes"*. It also happened regularly that the gesture was not understood despite being visible, and this led participants to repeat their movements several times before succeeding. The frequency of these errors usually decreased as they progressed through the experiment, which suggests that more training might have been required. This was accounted for by one of the participants: *"I, for sure, had a learning curve, but towards the end, I found it actually quite easy to manipulate her"*.

User engagement. Lastly, all participants reported that they enjoyed using the avatar. They unanimously said they were curious to know how well they scored. The comment boxes included positive feedback such as "*quite fun*", "*great*", "*very playful*", or "*strangely pleasant*" despite difficulties during manipulation accounted before.

5.2.6 Discussion

Ran with the Puppeteering mode of our proof-of-concept, this experiment shows that 3PP AR embodiment can successfully enhance the perception of physical space and estimation confidence:

(i) As expected, participants were more accurate when estimating their ability to act (moderate effect size) and the space occupied by their body (large effect size) with the help of their self-avatar

(ii) Participants were also more confident when performing mental projections of their body size (large effect size).

Therefore, these results allow validation of the approach we presented in Chapter 4 to enhance the perception of physical space. Additionally, they allow making sure that AR technology is reliable enough for such usage despite well-known issues linked to them.

The majority of participants found the Puppeteering mode to be useful, usable, and fun. We learned that the interaction technique we used to rotate the avatar needs to be improved as it was considered laborious by some participants. One participant suggested that it would be convenient to have a physical doll to put the avatar in the desired posture, as previously proposed in research in other contexts (Numaguchi et al., 2011). It could be interesting to study such a technique as it could allow for more direct and efficient enacting, as it does not rely on the detection of gestures that can be faulty.

Moreover, as the participants could not stand up and change their angle of view during the experiments, it was impossible for them to see parts of the avatar that were occluded by the avatar itself. Although none of the evaluation questions relied on this, it made it hard to tell when the avatar was in contact with surfaces behind. To remedy this problem, multimodal feedback could be provided (e.g. sound, visual cues, vibratory feedback) to inform the user when and where their avatar comes in contact with real surfaces that are occluded by other parts of its body. This could be useful in situations in which it is impossible to walk and change one's angle of view due to contextual or environmental restrictions.

5.3 Experiment 2

This second study aimed to further assess the strength of 3PP embodiment experiences by checking whether one's perception of real environments could also be improved **after** having controlled one's avatar, and not only while using it. To be able to compare its results with that of Experiment 1, we used the same system, experimental design, and evaluation as before.

Figure 5.7: Experimental environment on the beam. The environment on the grounded platform of Experiment 1 was also used in this experiment.

5.3.1 Apparatus

The apparatus of Experiment 2 was identical to Experiment 1, except that it had one supplementary environment: in addition to the arrangement on the platform in Figure 5.1, we also used the metallic beam suspended above this platform (see Figure 5.7).

5.3.2 Methodology

Experiment 2 was run in a between-subjects design as well. Its task, evaluation, and conditions were the same as in Experiment 1. The only difference is that the entire set of evaluation questions was asked at the end of the condition, after having performed all of the instructed actions.

- **Condition "R"**: the participants were asked to mentally visualize themselves performing the entire set of actions used in Experiment 1, one by one without interruption. Once all actions were mentally processed, the participants had to answer all the same questions as asked in Experiment 1.
- **Condition "AR"**: the participants had to implement this same set of actions with their avatar. After completing them all, they had to answer the same question too but without the help of their avatar.

In the AR condition, the HMD was removed from the participants before starting the evaluation so that they could no longer see their avatar when answering. The participants were not aware of the type of questions that they would get, and they were not specifically instructed to memorize what they saw while manipulating their avatar.

All participants went through both the R and AR conditions in counterbalanced order, but within different environments (see Section 5.3.1). The attribution of these environments was also distributed to either condition in a counterbalanced order. In other words, we had 4 groups which are illustrated in Figure 5.8:

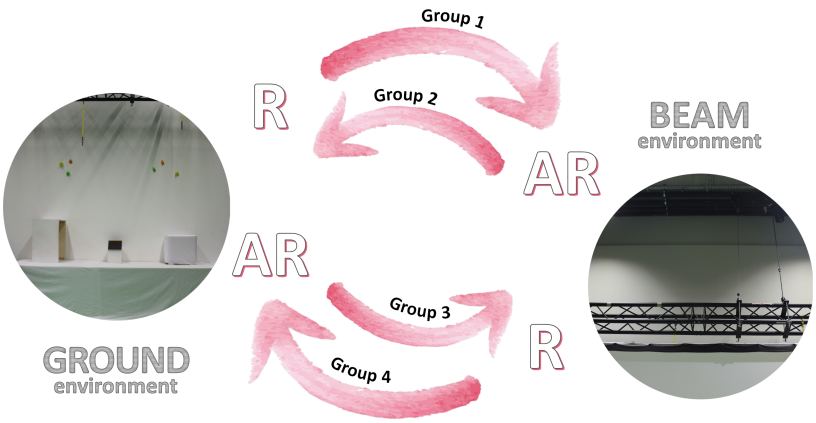


Figure 5.8: Experimental design of Experiment 2. The arrows represent the order in which the conditions were passed by the participants of each group.

The goal of this design was to analyze the effect of avatars in between subjects while having all participants test both conditions. After going through both conditions, the participants were invited to fill out a subjective questionnaire similar to the one in Experiment 1 (see Section 5.2.3). The hands-on time with the system was about 15 minutes and the whole session lasted approximately 1 hour.

5.3.3 *Measures*

The questions asked after each action to evaluate the participant's judgment abilities were exactly the same as in Experiment 1 (see 5.3.3). Figure 5.9 provides an example of an estimation-type question displayed on the beam. The post-experiment subjective questionnaire of Experiment 2 also contained the same comment boxes and questions as Experiment 1, except for item 7 which was reformulated, and for items 8 and 9 which were removed. Full details of this post-experiment questionnaire are available in Appendix A.1.3.

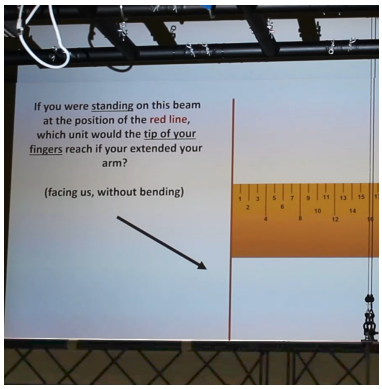


Figure 5.9: Example of estimation-type question that we asked, on the beam. The projected text says: "If you were standing on this beam at the position of the red line, which unit would the tip of your fingers reach if you extended your arm? (facing us, without bending)"

5.3.4 *Participants*

We recruited 16 participants from age 22 to 58 ($\overline{m} = 30.6, SD = 10.5$, 9 identified as males). Most had previous experience with AR headsets (12 of them). No compensation or course credits were issued, and all participants were unaware of the purpose of the experiments. Before the testing session, we collected the body measurements of the volunteers to generate their avatars.

5.3.5 *Results*

The same methodology was applied as in Experiment 1 to evaluate the accuracy of user judgments, their level of confidence, and their appreciation of the system (see Section 5.2.5).

Accuracy assessment

Shapiro-Wilk tests show that the scores did not follow a normal distribution, so we used non-parametric tests to evaluate the significance of differences.

Descriptive analysis. Generally speaking, participants seem to have underestimated their body sizes in both R and AR conditions, regardless of the environment (beam or ground). The results of our descriptive analysis are summarized in the boxplots of Figure 5.10.

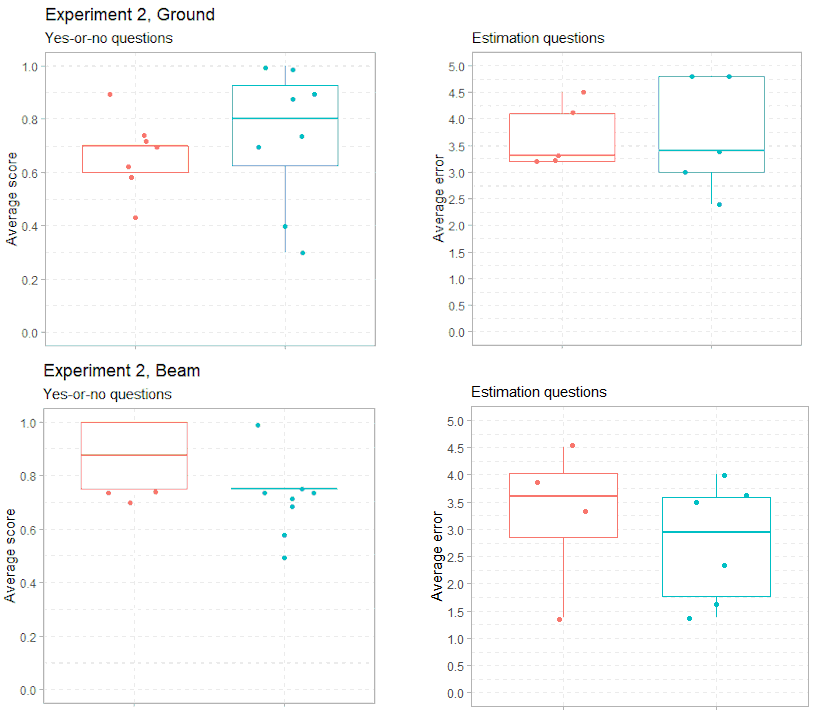


Figure 5.10: Boxplots representing scores of the yes-or-no and estimation questions asked during Experiment 2.

Statistical tests. Our statistical analysis did not show clear evidence that having used the avatar led participants to have a more accurate perception after removing the AR headset for either environment (Wilcoxon rank sum tests, yes-or-no questions: $p_{ground} = 0.34$, $p_{beam} = 0.19$; estimation questions: $p_{ground} = 0.69$, $p_{beam} = 1$). Further study will be needed to determine if manipulating a self-avatar in unexplored distant locations allows for assimilating an experience that can be used from memory.

Confidence level

Shapiro-Wilk tests show that the confidence data of Experiment 2 does not follow a normal distribution. We still used non-parametric tests for both as the data is not continuous and the number of participants is rather small.

During the evaluation, participants frequently accounted for the difficulty of the questions and communicated they were very unsure of their answers. Despite subjective feedback suggesting that they were more confident in the AR condition, we found no significant difference in the confidence and error margins between the R and AR conditions (Wilcoxon rank sum tests, yes-or-no questions: $p_{ground} = 1$, $p_{beam} = 0.81$; estimation questions: $p_{ground} = 0.40$, $p_{beam} = 0.30$).

User feedback

Figure 5.11 shows the results of the post-experiment questionnaire of Experiment 2. The questionnaire was nearly the same as in Experiment 1, except for item 7 which was reformulated, and for items 8 and 9 which were removed to shorten the content.

Help provided by the avatar. As in Experiment 1, the subjective feedback shows that 80% of the participants found the avatar helpful despite not seeing it during the evaluation. Several participants of this experiment mentioned that the avatar had allowed them to correct their perception of distance or size: *"I saw myself much smaller, [the avatar] allowed me to raise my estimates"*, *"I could better realize the relative position of each object. For example, ah, the distance between the chain and the block is not that big"*.

Trust in the avatar. In this experiment, the participants were not explicitly instructed to memorize the space that their avatar occupied and the places it could reach while manipulating it. Nevertheless, they often commented that they were trying to recall what they had observed and relied on their memory to answer. For example, one participant wrote: "*I used what I saw with the avatar to estimate the answers to some questions*". The participants regularly expressed that they wished their avatar was still visible as it was sometimes hard for them to remember what they had seen. We did not witness signs of mistrust towards the avatar like in Experiment 1.

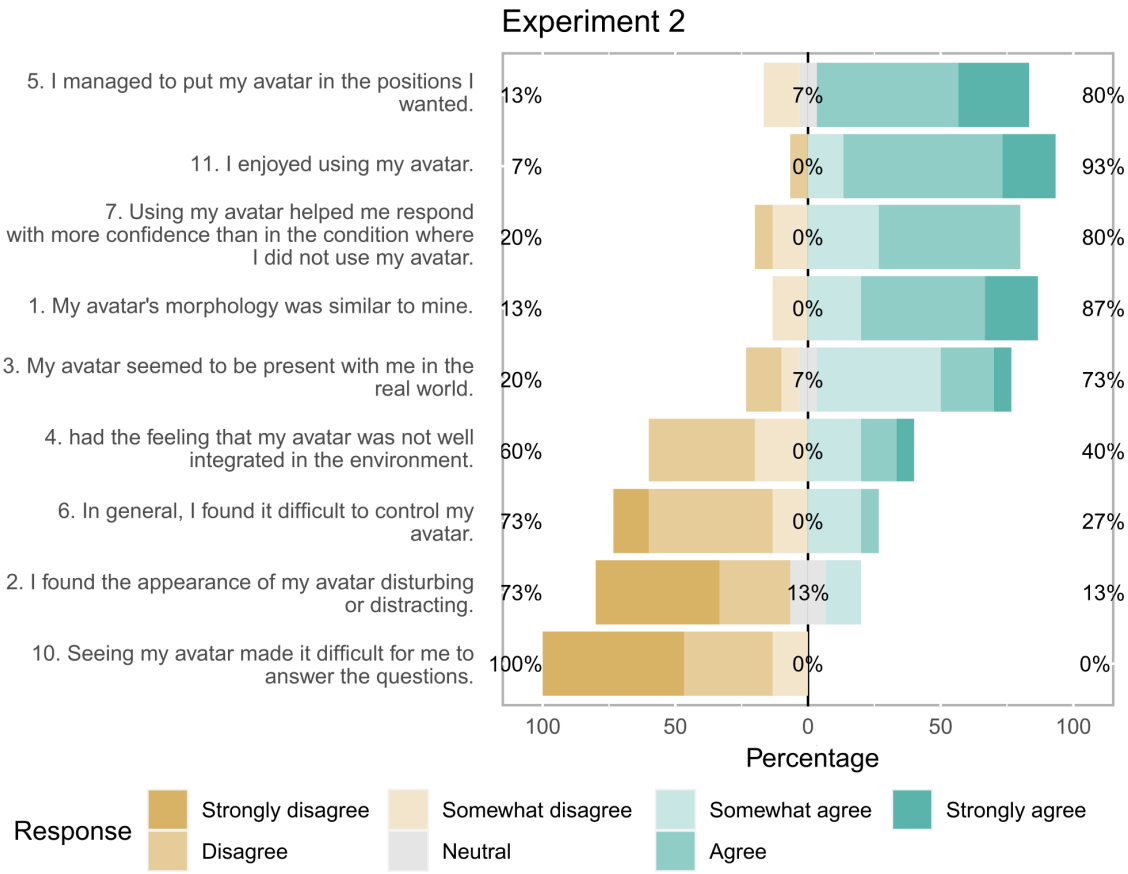


Figure 5.11: Results of the subjective post-experiment questionnaire of Experiment 2 (translated from French). Detailed percentages are available in Appendix A.2.3

Avatar appearance. Like in Experiment 1, most participants agreed that the avatar resembled their own body (item 1). Among the two that did not agree, one participant observed that the avatar's distribution of fat was quite different from their own. This is probably linked to the limited number of input parameters of Virtual Caliper, which does not include traits like muscularity. Similarly, most participants did not find their avatar's appearance to be disturbing or distracting (all except 2). Among those that did not like it, one mentioned that the idle animation of the avatar made them uncomfortable.

3D registration. As in Experiment 1, some participants experienced environment detection issues that led them to have more mitigated answers. These issues occurred when the spatial mesh built by the Hololens 2 was updated inaccurately, which happened more often when the AR condition was performed on the beam due to its height and angle of view. As a result, the avatar appeared occluded by a virtual wall or residual artifacts. Item 3 reporting that the avatar felt "present" with them in the real world had high scores (73%) regardless of that.

Avatar control. We had similar comments and ratings as in Experiment 1 on items 5 and 6, suggesting that the users were satisfied with the avatar's control. Room for improvement was pointed out regarding the rotation of the avatar, found difficult by several. Another participant suggested including more feedback to better perceive when the avatar is in contact with real surfaces. Gesture detection issues were also experienced by users, especially when they were unfamiliar with the headset[3].

[3] As we wanted to preserve the comparability of this experiment with the previous one, we had decided not to lengthen the user training as suggested by the subjective feedback we collected in Experiment 1.

User engagement. All participants except one enjoyed interacting with their avatars. The only one that disliked it had done the AR condition in the beam environment and commented the following: "*It is more the location of the avatar (far and high) that is a pain rather than the control of the avatar itself*". This feeling was shared by several others who reported they found it hard to select the spheres depending on their angle of view and that this difficulty slowed them down: "*It is quite complicated to control the avatar when it is very high or low*". We expect that the next generation of AR technology and improvements in the proposed interaction techniques may solve the usability problems that they described.

5.3.6 Discussion

Through this experiment, we wanted to verify whether a self-avatar could go beyond being a visual reference and provide a way to gain experience that can be used after manipulating it. However, we could not find clear evidence that the avatar helped users improve their representation of physical space as we found no significant difference between the confidence and accuracy of their estimations, with and without having used their avatars.

The design of our experiment has probably played a role in these results: to avoid a transfer effect, participants manipulated their avatars without knowing the questions we would ask them. It seems that this led them to focus their entire attention on how to achieve the posture we were describing, and not on how the avatar was fitting in the environment.

Such behavior was described as "attentional tunneling" in previous work and seems to be a common issue in AR studies (Syiem et al., 2021). The research of Fischer et al. (1980) was one of the earliest to document it. Their study compared the performance of pilots in a cockpit simulator when displaying flight information (altitude, path, etc.) on a Head-down Display (HDD) to when using a Head-up Display (HUD). Although the performance of the participants was higher when using the HUD, pilots tended to miss more unexpected occurrences in the outside world compared to when using the HDD. This result was unexpected as the HUD retained the outside world in foveal vision whereas the HDD forced users to move their sight between two contexts.

According to Syiem et al. (2021), it is not the virtual content that is responsible for attentional tunneling effects, but the type of task: through two studies, they showed that participants tended to excessively focus on virtual content when they were asked to complete a task involving this content, but not when this task was suppressed. Such focus on a particular content — whether it be real or virtual — can lead to inattentional blindness. This phenomenon refers to the failure to notice fully-visible content when attention is diverted elsewhere. A famous illustration of this effect can be found in the study of Simons et al. (1999) where participants fail to see a person passing through the scene in a gorilla suit as they try to count the number of passes made by a basketball team.

We believe that a similar kind of phenomenon occurred with our participants. Some participants accounted for this: "*I was more focused on manipulating the avatar than on its size in space*". As they were not expecting to be asked questions related to the physical environment in which they interacted, it seems that they selectively looked at the avatar and did not consciously perceive the objects around it despite their proximity. Consequently, they could not memorize and remember the information that their avatar was meant to show when answering our questions. Additionally, it is possible that the high degree of manipulation required by the Puppeteering mode further amplified these effects. The participants needed to perform multiple actions on their avatars to achieve the postures we asked them to implement, which probably increased further their engagement.

To conclude, our experimental design may have impacted the results we observed. In Experiment 3, we rethought our protocol to include user priming, letting them know that they should use their avatar to interpret the affordances of the environment. We additionally reduced the task to estimating only one affordance (the passability of a door) instead of a sequence of different ones. Lastly, we replaced the Puppeteering mode with the Physical Control mode to simplify the control of the avatar.

5.4 Experiment 3

This experiment was designed and conducted in collaboration with Prof. Bernard N'kaoua and Léa Migliore as a part of her Master 2 internship.

After observing the results of Experiment 2, we decided to conduct Experiment 3 with a new methodology to see whether our experimental design was responsible for the absence of the effects we were expecting. Therefore, the purpose of Experiment 3 was the same as Experiment 2: we wanted to verify whether having used an avatar to simulate an action could help users during their future estimations when the AR embodiment system is no longer available.

This experiment's protocol was inspired by Bhargava et al. (2020) who compared the passability judgments of participants in VR with those made in the real world. The task consisted in judging whether

it is possible to pass through a door without turning one's shoulders while standing a few meters ahead. The participants repeated this task with 13 door openings presented randomly, ranging from 0.7 to 1.3 times the size of their shoulders with an increment of 0.05.

Compared to the original experiment of Bhargava et al. (2020), we used AR instead of VR and displayed an avatar in the 3PP instead of the 1PP. Since our objective was to investigate the resilient effects of avatar manipulation, we additionally asked the participants to make their judgments **after** having let them test the door's passability with their avatars, rather than while seeing it. The experiment was divided into two phases:

- **Preparation**: First, the user could test the passability of the 13 doorway sizes by making their self-avatars walk through them.
- **Evaluation**: Then, the avatar was hidden and the same doorway sizes were presented again. They had to judge each doorway's passability and we recorded their answers.

Unlike in Experiment 2, the participants were aware of these two phases in advance. We explicitly asked them to memorize each doorway's passability while using their avatar as we would evaluate them with the same doorway sizes in the evaluation phase. This allowed taking the experiment closer to ecological settings since users should normally use such a system with a well-defined goal, knowing the type of information they want to obtain.

Additionally, we thought it would be interesting to check if the user's performance changed when judging the passability of a real door compared to a virtual one displayed in AR. Indeed, as previous experiments showed perceptual issues linked to AR, we wanted to verify their impact on the judgment of affordances. Therefore, we performed a cross-evaluation of the following two factors:

$$\textbf{Door virtuality}(\textit{real}/\textit{virtual}) \times \textbf{Prior use of the avatar}(\textit{yes}/\textit{no})$$

We hypothesized that the participants in the real door conditions would perform better than those in the virtual door conditions and that those who had used their avatars would improve their estimation performance in both cases.

5.4.1 *Apparatus*

System. Like in the previous experiments, we used a Microsoft Hololens 2 to display the participant's avatar. The system was the same as before, except that we replaced the Puppeteering mode with the Physical Control mode. We made this change to reduce the focus that the participants would put on their avatar's control as they wouldn't have to handle its posture in this mode. Indeed, in this mode, the animation system automatically puts the avatar into the posture upon pressing the gamepad's buttons.

The participants were given an XBOX Series X/S Bluetooth controller connected to the headset to move the avatar and record their

answers with the A and B buttons (for "Yes" or "No" resp.). All commands except walking were disabled to avoid variability in how the users controlled their avatars.

The 13 doorway sizes were computed with a C# script upon entering the user's shoulder size through a UI on the Hololens (Figure 5.12). Like in the experiment of Bhargava et al. (2020), they ranged from 70% to 130% the participant's shoulder size with 5% increments. These sizes were then used to create two lists of doorways to be shown during the preparation and evaluation (in random orders). The preparation list contained each of the 13 doorway sizes once whereas the evaluation list featured them 5 times each. To consult these lists without having to wear the headset, the experimenter could import it onto a laptop over WiFi through the Windows Device Portal.

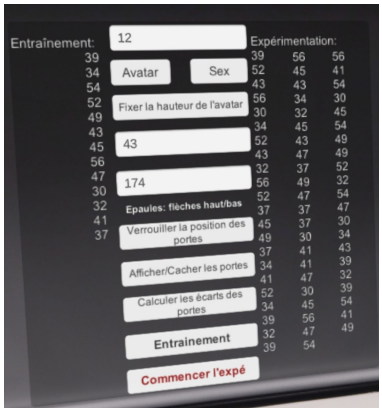


Figure 5.12: Configuration UI used to show/hide the virtual door and avatar, to configure the avatar, and compute the doorway sizes for each participant. The list on the left is for the preparation phase, whereas the one on the right is for the evaluation phase.

Avatar. To generate the avatar, we did not use Virtual Caliper as before because it did not allow controlling the avatar's shoulder size. We used the UMA 2 Unity package instead, like in the paper of Bhargava et al. (2020). UMA allows personalizing the features of 3D mesh models within the runtime. The avatar's shoulder size, height, and gender were configured on the Hololens through a Bluetooth keyboard connected to it (see Section 5.4.2 for details). Other body parts were not changed. Lastly, we covered the avatar with the same blue abstract texture as previously (see Figure 5.13).

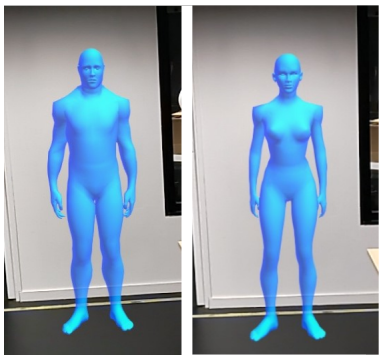


Figure 5.13: *Left* Male avatar. *Right* Female avatar.

Environment. The experiment took place in a room equipped with a pocket door of 2m x 1.6m, manually moved by the experimenter in conditions RD and RDA (see Figure 5.14, left). For conditions VDA and VD, we modeled a 3D life-sized replica of this door, which the experimenter set in the exact place of the real one (see Figure 5.14, right). The virtual door was moved through a C# script. The participants stood 2 meters ahead of the door to judge its passability. This distance allowed the user's field of view of the door to be large enough to see the entire door in all conditions.

Figure 5.14: *Left* Real door. When fully opened, it completely recedes into the wall frame. *Right* Virtual door. Modeled with Unity primitive objects.

5.4.2 *Methodology*

The experiment lasted about 45 minutes. It was run in a between-subjects design with the following four conditions:

- **Condition "RDA" (real door after avatar control)**: wearing the AR headset, the participants had to walk their avatar through the physical door ahead of them and orally answer "Yes" or "No" to whether they could pass without turning their shoulders (**preparation**). They repeated this for the 13 door openings presented sequentially. Then, they had to remove the AR headset and repeat these judgments with these same door openings, each presented 5 times in a random order for a total of 65 trials (**evaluation**).

- **Condition "RD" (real door)**: the participants didn't train themselves with their avatars and would directly go through the 65-trial evaluation of the physical door's passability.
- **Condition "VDA" (virtual door after avatar control)**: same as condition RDA, but with the virtual door.
- **Condition "VD" (virtual door)**: same as condition RD, but with the virtual door.

Configuration. Upon the arrival of the participant, the experimenter asked for their body height and measured their shoulder size with sewing tape. The latter was collected to compute the doorway size list. The participant then filled out Questionnaire 1, collecting demographic data and evaluating their relationship to AR and VR technologies. In the meantime, the experimenter put on the AR headset, launched the experiment's app, and configured the condition to be passed through the UI in Figure 5.12. For conditions VDA and VD (with a virtual door), they fully opened the physical door and instantiated the virtual door in the room. Then, they placed this door in the frame of the real door and locked its position. In conditions RDA and VDA (with an avatar), the experimenter then instantiated a UMA generic avatar through the UI. As the Hololens 2 floor detection is not very precise, the experimenter manually adjusted the avatar's position on the ground. Once it appeared to be standing on its surface, they locked the avatar's movements on the vertical axis and adapted its gender to match the user. To adjust the avatar's height and shoulder size, they asked the participant to stand in front of them and superimposed the avatar onto their body. Then, they adjusted the sizes with the Bluetooth keyboard connected to the headset until it visually matched the user[4].

[4] UMA does not support entering cm values to customize the avatar. Instead, it provides sliders ranging from 0 to 1 to modify each feature. We adopted this method to customize the avatar for this reason.

The conditions requiring AR (RDA, VDA, and VD) were preceded by an eye calibration. In the conditions involving an avatar, the participants were introduced to it and asked to compare its shoulder size and body height with their own. Before starting, the participants were invited to take the gamepad and stand behind a tape line on the floor, located 2 meters ahead of the door. They were then asked to close their eyes while the experimenter went to the other side of the door and hid behind the wall. Then, the experiment could begin.

Preparation. In the conditions using the real door, the experimenter would consult the preparation phase's doorway list on their laptop to slide the door into the next position. To do so, they first slid the door back and forth thrice before sliding it to the actual width. They used a measuring tape attached to the top of the door frame (out of the participant's sight) to set the door to the right distance. Then, the participant was asked to walk their avatar back and forth through the door and to answer whether they they recorded the participant's oral answers on a spreadsheet.

In the conditions using the virtual door, the participant recorded their answers through the gamepad. Pressing the A or B buttons

would slide the virtual door automatically into the next position in the list, so the experimenter didn't have to handle that either.

Evaluation. The same process was repeated as in the preparation phase without the avatar and with the evaluation phase's doorway list. At the end of the 65 trials, equipment was removed from the participant and they were invited to answer Questionnaire 2 evaluating their experience.

5.4.3 *Measures*

The first part of Questionnaire 1 collected demographic data to characterize the sample (gender, age, previous VR and AR experience, and video game play time per week). The second part assessed the participant's wariness of technology over 5 items, issued from the work of Anania (2019). They were meant to identify potential bias related to personal opinion and confidence in the equipment. We decided to include this scale upon observing mistrust in the AR avatar from some participants of Experiment 1. Full details are available in A.5.

During their task, the participants were asked to answer one single yes-or-no type question: "Do you think you could pass the door without turning your shoulders or curling up?". We recorded their answers for each doorway size presented during the evaluation phase. They were not given feedback about the correctness of their answers. Then, we computed the three following scores to assess their accuracy and the type of errors they made:

- **Accuracy score**: number of correct answers on the 65 trials.
- **Underestimation score**: number of times the participant wrongly estimated that they could not pass.
- **Overestimation score**: number of times the participant wrongly estimated that they could pass.

Questionnaire 2 contained 14 items on a 5-point Likert scale, and 5 comment boxes letting the participants write their thoughts on the AR system, virtual door perception, avatar's control and appearance, help provided by the avatar, and their global appreciation. Full details are available in Appendix A.4.

5.4.4 *Participants*

We recruited 42 participants from age 19 to 46 ($\overline{m} = 28, SD = 6.81$, 26 identified as males). Most had previous experience with VR or AR headsets (35 of them). No compensation was issued, and all participants were unaware of the purpose of the experiment.

5.4.5 *Results*

After verifying that the variables were homogeneous across groups (i.e. age, relationship to technology), we ran a descriptive analysis and statistical tests on the three accuracy scores. Shapiro-Wilk indicated that all the data met the assumptions for parametric tests. The homogeneity of variances was also validated by Levene's test.

Accuracy assessment

Descriptive analysis. From our descriptive analysis, it is not clear how the virtuality of the door and prior use of the avatar influenced the user's accuracy. In the virtual door conditions, the average accuracy score was roughly the same with and without prior use of the avatar. In the real door conditions, scores seem to have varied in a wider range when the avatar was used. Table 5.1 summarizes the scores for each condition.

Table 5.1: Descriptive statistics on the accuracy scores.

	Door	Avatar	N	Mean	SD	Min	Max
Score /65	Virtual	With	11	55.8	4.40	48	62
		Without	10	55.5	6.13	44	62
	Real	With	11	52.5	8.90	37	61
		Without	10	57.0	5.03	45	62

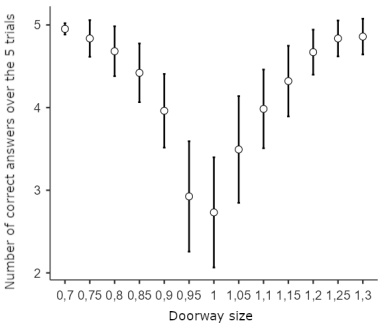


Figure 5.15: Average success score for each of the 5 trials of each doorway size.

As in the experiment of Bhargava et al. (2020), the passability ratio (i.e. doorway width divided by shoulder width) predicted the user's judgments: participants were more likely to consider a provided doorway passable as its width increased, and impassable as its width decreased (see Figure 5.15).

Statistical tests. We performed a two-factor ANOVA to compare the global accuracy scores of the participants over four conditions. The test did not show a significant difference for either factor (avatar: $p = .30$, door: $p = .65$). The interaction between these two factors was not significant either (interaction: $p = .23$). These results suggest that there was a relatively homogeneous average success between the groups. Therefore, we could not validate our hypothesis that the prior usage of the avatar would globally help users to improve their accuracy when judging a door's passability.

To compare the type of error made in each condition, we performed ANCOVAs that included the overestimation and underestimation scores. We found a significant difference in the underestimation scores of the users who had used their avatars compared to those who had not ($F(1,37) = 51.56, p < .001$***). There were fewer underestimation errors when the avatar was used, for both the real and virtual doors. Therefore, it seems that using the avatar generally helped make fewer underestimation errors. We could not find significant results on the overestimation scores.

To complement these tests, we verified whether the accuracy of the participants evolved over time through a within-subjects analysis. From repeated measures ANOVAs, we found the following results:

- **Global accuracy**: The user's accuracy significantly deteriorates over time, regardless of the condition ($F(152) = 4.57, p = .017$*). This deterioration was faster when the avatar was used ($F(152) = 2.78, p = .029$*).
- **Overestimation**: The overestimation errors significantly increase over time, regardless of the condition ($F(152) = 4.57, p = .002$**). During the first trials, however, there were slightly fewer overestimation errors when the avatar was used ($F(152) = 6.13, p < .001$***). The overestimation score joins that of the other conditions after the fourth trial.
- **Underestimation**: There were fewer underestimation errors during the trials with the virtual door than with the real door ($F(152) = 1.96, p = .02$*). The avatar did not significantly impact the evolution over time of the underestimation errors.

We found no significant interaction effects between the avatar and door factors for the global accuracy, overestimation, and underestimation scores.

Lastly, we ran an ANCOVA to test the effects of demographic characteristics of users on these results. The variables tested were average video game play time per week, previous VR and AR experience, age, gender, as well as the relationship to technology score. We also ran Pearson correlation tests. We could not observe any significant results.

User feedback

Figure 5.16 shows the results of the post-experiment questionnaire of Experiment 3. From the oral feedback and comment boxes, it seems that most participants found the task difficult, regardless of the condition they passed. This was confirmed by item 9 on the task's easiness, to which 91% of participants disagreed.

Help provided by the avatar. The subjective feedback shows that 86% of the participants of RDA and VDA conditions used and relied on their experience and memory of the avatar to answer the evaluation questions (item 11). Most of them gave high rates to items 10 and 12 (73% and 77% resp.), showing that they found that this preparation was quite helpful to make their judgments and allowed being more confident when answering. This is also clearly reflected in some comment boxes. One of them detailed the following aspects:

> "The avatar allowed me to really project myself. Without it, I would never have known if I could pass or not and would have tended to answer randomly. I was surprised during the training phase to see that I could pass at times, which allowed me to adapt my answers during the test phase."

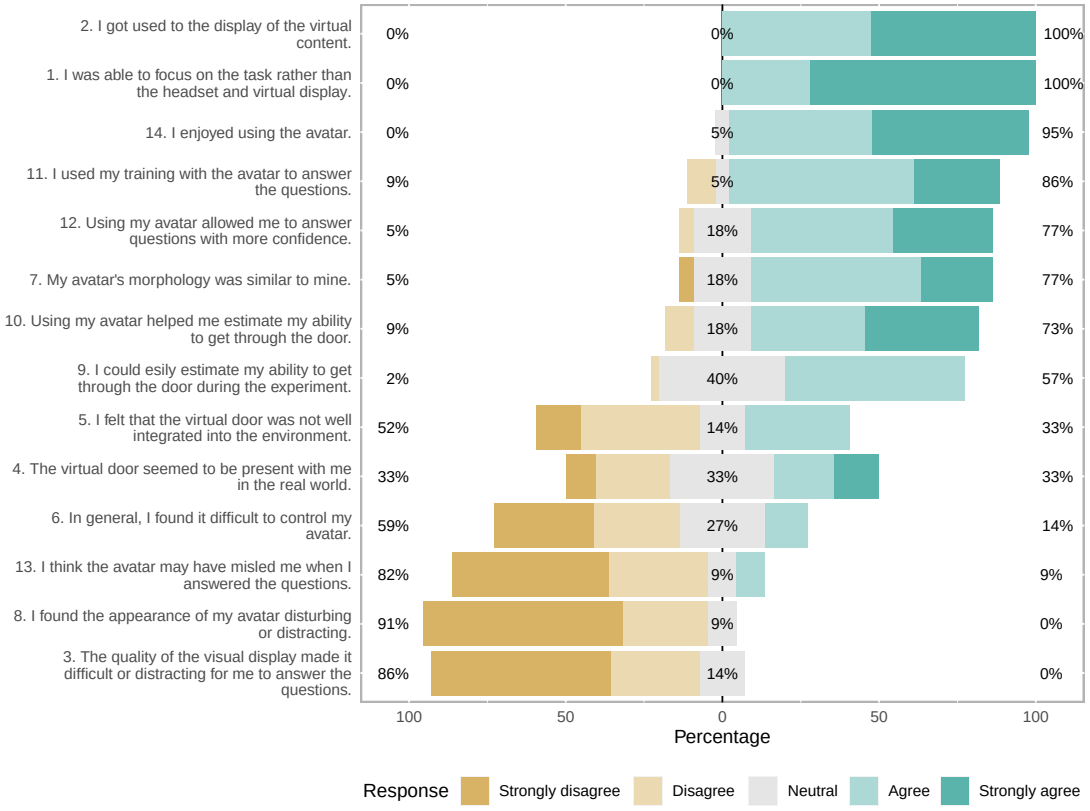


Figure 5.16: Results of the subjective post-experiment questionnaire of Experiment 3 (translated from French). Detailed percentages are available in Appendix A.4.

Other comments included *"I think that without the avatar I would have underestimated the width of my shoulders much more"* and *"It helped me to situate myself much better in space"*. These multiple feedback contrast with the lack of significant effects of our analysis. It seems that even though most participants found their avatar useful, their performance could not fully benefit from it. One participant wrote a potential explanation for this: *"The memory of the avatar fades as the experiment progresses. At first it helps, then you forget"*. We discuss this in Section 5.4.6.

Avatar perception. The UMA avatar seems to have been well accepted by the participants despite being much less personalized than in previous experiments: 77% answered that they found that their avatar correctly represented their ability to pass the door (item 7). No participant found the avatar's appearance disturbing or distracting (item 8), and most do not think their avatar misled them during the task (item 13). Questions on the control of the avatar returned more mixed feelings (item 6). Whereas some found it easy to use (59%), others commented that it was difficult to put it exactly where they wanted because of the sensitivity of the control. One participant recommended adding a lateral control function that doesn't rotate the avatar.

Door perception. Items 4 and 5 show that the virtual door was not always perceived as perfectly integrated into the real environment. Whereas some provided positive comments (e.g. *"sometimes you can confuse shapes without knowing if the line is in the real or in the virtual"*), others were very sensitive to the AR rendering and not convinced by it (e.g. *"The virtual door was moving a little too much compared to the small movements of my head"*).

User focus. All participants reported that they could focus on their task and were not particularly distracted by the virtual content, but eventually got used to it (items 1, 2, and 3). This suggests that the attentional tunneling issues we observed in Experiment 2 were corrected in this experiment.

User engagement. All participants enjoyed doing the task, despite it being repetitive (item 14). The participants often left their contact information to be informed of the results and to be invited to future experiments.

5.4.6 Discussion

We were interested in whether the memory acquired while manipulating the avatar is assimilated and may be used in analogous situations. The feedback collected from the RDA and VDA conditions shows that users felt the preparation with their avatars was useful to make their judgments. However, we could not find evidence that this preparation had actually helped them as there was no significant difference with the other conditions.

Although the evaluation started immediately after the preparation phase, it seems that the users were not able to remember the different doorways' passability they had tested with their avatars. One subject commented the following: *"After having forgotten a bit for which gap it was possible to pass or not, the task became a little too hard and I was obliged to answer by instinct"*. Other participants shared similar impressions orally during the experiment.

We believe that this impossibility to recall what they saw during the preparation phase was experienced by most participants. Indeed, the task required memorizing the door's passability with a centimeter-level accuracy, which is difficult in normal circumstances too. We were expecting that the participants would improve their judgments over time in the RD condition at least, as in the experiment of Bhargava et al. (2020). In their real door condition, the authors found that the user's probability of answering correctly was increased by 3% after each trial, despite not being given feedback on their answer's correctness. They could not observe such an improvement in their VR condition. We hypothesized that having used the avatar during the preparation phase would amplify the improvement they observed in real-world conditions. However, we could

not observe any improvement — with or without the avatar — even though our protocol followed the same methodology as theirs. The reason for this remains unclear at this point and will need further investigation.

Lastly, we could find no difference between the real and virtual door conditions in terms of estimated chances of passing through them. This could imply that the information that guides the affordance perception of the virtual door accurately reflects the information that is present in the physical door. However, the collected post-experiment feedback shows that one-third of users felt that the virtual door was not correctly integrated into the real environment. One possible way to interpret this is that the sense of the virtual door's presence is not linked to affordance perception related to it. In other words, the user's judgment accuracy may be independent of how realistically the door is integrated.

5.5 Conclusion

In this chapter, we explored how self-avatars and MR can be used to extend our perception of physical environments by expanding our range of action to areas outside of our peripersonal space. We used two control techniques of the AR embodiment system designed in Chapter 4 to evaluate the validity of our approach through three exploratory experiments. Our results highlight how such use of avatars has the power to improve the user's understanding of their options during decision-making.

However, the advantages it provides seem limited in time: once taking the AR headset off, it is hard to recall what one saw while manipulating the avatar after a short period. Therefore, the system we present seems to be most interesting when the user can use it while in action.

Finally, from the experiments we ran, we also learned that the Puppeteering mode requires cognitive efforts that can generate attentional tunneling effects. We recommend using it when the number of iterations required to achieve a posture is low, or to use other modes otherwise (e.g. Physical Control, as in Experiment 2).

Future work

The experiments we presented involved two of the control modes we implemented in Chapter 4. As a follow-up, future work should assess the usability of all three control modes by measuring time performance. Embodied cognition being linked to the SoE, measuring the user's feeling of ownership and agency towards the avatar may also reveal valuable information. We did not include a measure of these senses as our goal was to primarily confirm that 3PP avatars in AR could improve affordance perception, but we encourage future experiments to investigate them in dedicated studies.

Secondly, it would be interesting to reproduce the experiments we presented with an extra condition where the participant actually performs the actions asked to be judged. This would allow comparing more directly the experience acquired in the third-person through one's avatar with the one acquired by directly acting. Additionally, Medeiros et al. (2018) previously identified that the avatar's appearance could impact spatial awareness in VR, but this fact was not studied in AR. We would therefore recommend pursuing research on this effect.

To finish, AR displays still have technical limitations that impede exploiting avatar embodiment in physical space to its full potential. In particular, contact identification between real and virtual surfaces is a well-known perceptual issue of XR interfaces. This is partly because rendering light and shadows on virtual objects in real time is still a hard problem. Without shadows, holograms seem to float in mid-air instead of resting against surfaces. This is an ongoing problem, and we did not evaluate the impact of such an issue on physical affordance perception. It will need to be investigated in future work.

Reality is merely an illusion,
albeit a very persistent one.

Albert Einstein

6

Transitioning from "being oneself" to "being one's avatar"

Work in progress

This chapter details the collaborative work done with Prof. Kiyoshi Kiyokawa at the CARE laboratory throughout a two-month stay at the Nara Institute of Science and Technology (NAIST). Our goal was to help users build a stronger SoE toward their avatars. To do so, we explored ways to mentally prepare users to embody their avatars. We conducted a pilot study in VR to test a preparation protocol applied to embody an avatar representing the Hulk from the Marvel franchise. The protocol we designed involves a mindfulness exercise and an acting activity to help participants imagine that they are Hulk before seeing themselves in its body. We hypothesized that this preparation process would help users to build a stronger SoE. As the results of this collaboration were still being processed when writing this thesis, only part of them are presented. Lastly, we briefly mention the second project in which we participated, exploring the usage of graphical transitions to smooth avatarization experiences. This project was led by Riku Otono of the CARE lab.

Contents

According to Harrell et al. (2017), the Avatar Dream stands on two components: a technical one and an experiential one. The first one aims at allowing the control of avatars as naturally as one's body, whereas the second one aims at making users live lifelike experiences that would otherwise be impossible. While research has made much progress on the first technical component, it seems that the creation of believable embodiment illusions still often fails. In fact, it is common to observe avatars being rejected by their users even with state-of-the-art systems.

This difficulty is often explained by the existence of subjective factors (Dewez et al., 2019). Depending on the user, they can lead to either a propensity to BOIs or reluctance to feel embodiment. Therefore, it cannot be guaranteed that an embodiment system will generate the same experience across users. In addition to making experimental research on BOIs complicated, this variability also makes it impossible for some users to benefit from their effects. Indeed, "*virtual embodiment of a relevant avatar is not, in itself, sufficient to induce a Proteus Effect*" as Ash (2016) wrote.

While removing subjective factors is impossible, strategies can be elaborated to reduce the variability that they cause. We explored two of these strategies over a collaboration with colleagues[1] of the CARE lab:

- **Mental preparation**: consists of mentally conditioning the user to embody their avatars before the body avatarization starts.
- **Graphical transitions**: consists of smoothing the experience of the switch between the real and virtual bodies by applying visual effects.

[1] Riku Otono, Monica Perusquía-Hernández, Naoya Isoyama, Hideaki Uchiyama, and Kiyoshi Kiyokawa participated in this collaboration project.

This chapter mostly details the research on the first strategy (mental preparation). The second one will be mentioned more succinctly as it was principally investigated by the master student Riku Otono and his supervisors. The initial results of this collaboration suggest that such approaches are possible to improve the strength of BOIs. However, the research on both of these subjects (ran in parallel) was still in progress while writing this manuscript. As such, this chapter will only provide preliminary results.

6.1 Mentally preparing users

Embodiment experiences can be divided into three main phases: *pre-embodiment*, *embodiment*, and *post-embodiment* (see Figure 6.1). In most experiments on the SoE, the stimulation to induce the BOI is typically applied after entering the embodiment phase, once the user can already experience their avatarization. Then, the post-embodiment phase is often used to assess the resilient effects of the avatar on the users or to collect additional subjective data.

The pre-embodiment phase, on the other hand, is rarely exploited: during that time, participants are usually told about their task and

confronted with their avatars without further due. Users have to directly adapt to their new bodies and learn what it affords. To make this change less unsettling, we propose to familiarize them with their avatars beforehand and to mentally prepare them to embody them.

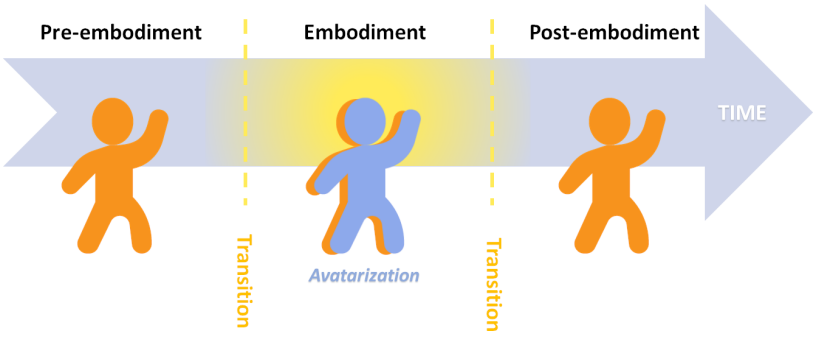


Figure 6.1: Main phases of embodiment experiences. *Pre-embodiment* Before the user embodies their avatar. *Embodiment* Time during which the user experiences and controls their avatar. *Post-embodiment* Time after the embodiment ended.

6.1.1 *Related work*

VR research indicates that personality traits (Dewez et al., 2019), hypnotic suggestibility (Walsh et al., 2015), and sensory suggestibility (Marotta et al., 2016) are determinants of BOIs[2]. As their relationship with the SoE is not well understood, very few papers explored ways to reduce their effects. We only found one article discussing how the mental conditioning users could help induce stronger BOIs: in their paper, Walsh et al. (2015) suggested future work to try using hypnosis to manipulate embodiment perception. According to them, putting users in a hypnotic state would enhance their attentional focus on the stimulus applied (e.g. VT stimulation) and therefore allow for a stronger response to it. However, performing hypnotic procedures requires training and their success remains dependent on the user's hypnotizability (Cardeña et al., 2014).

[2] The dominance of such top-down factors has not yet been discussed in AR to our knowledge, but it seems likely that they also exist in such contexts.

As mentioned before, previous work showed that avatar embodiment can be used to create empathy towards other people and influence their behavior (Yoon et al., 2014). In our research, we want to verify if the opposite is also true: does building empathy and deliberately changing one's behavior allow for a stronger embodiment of avatars? In particular, we investigated the use of activities such as mindfulness and role-playing which were shown particularly efficient to build empathy towards others (Bertrand et al., 2018).

Mindfulness exercises often require subjects to concentrate on their breathing and let thoughts run without trying to refrain from them. They can be oriented to create perspective-taking experiences and enhance empathy-related abilities by guiding the user's imagination (Hildebrandt et al., 2017). According to research by Lueke et al. (2015), practicing mindfulness for just 10 minutes can help people lessen their prejudices against other groups. Acting is also an efficient way to foster perspective-taking abilities (Whitebread et al., 2012). Role-playing techniques have been applied in a variety of settings, including therapy and conflict resolution (Bertrand et al., 2018).

To create multisensory experiences, they combine physical dynamics (e.g. switching seats with another participant) and story-telling. The use of such practices to help users accept avatars as their own bodies was never investigated to our knowledge. We present research on this topic for the first time.

6.1.2 *Design of a mental preparation protocol*

We propose to combine mindfulness and acting exercises over a 10-minute protocol aiming to guide users into imagining they have become their avatars. This protocol is meant to be used with avatars that represent specific characters bearing a personality or stereotypes that can be imitated. Therefore, it is not compatible with abstract avatarization forms or with avatars that are not different from the users. Readers should however note that the experiment we present does not aim to cover every case, but acts as an entry point to investigate how to better exploit pre-embodiment time.

First, with their eyes closed, the participant listens to the experimenter's instructions asking them to focus on their bodily signal (breath, heartbeat, skin sensations). The experimenter also describes the emotional states relevant to the avatar and asks them to imagine it. The participant is then asked to picture their bodies progressively turning into their avatars as described by the experimenter. After that, the user stands up and tries to act like the character represented by their avatar. As in the previous part, the experimenter provides indications on how to impersonate their avatar. The participant is then finally immersed into their virtual body in the 1PP.

To prevent the user from imagining something too different from what the embodiment system will display, we propose to introduce them to their virtual bodies before this process. This can be done by simply showing the avatar's 3D model on a screen and explaining that this is what they should be picturing. Secondly, it seems necessary to brief the user on who the avatar is before starting. Otherwise, it might be difficult for them to imagine the overall behavior they are expected to adopt. Therefore, the experimenter provides background information about the avatar before starting. They can then check the participant understood what they should imagine by asking them to describe the avatar they are to embody.

This entire process aims to make the user a full-fledged actor of their BOI by making them participate in its construction. As a consequence, we hope to generate mental conditioning that will facilitate the acceptance of the avatar once embodying it. The protocol described above can be summarized in the following steps, illustrated in Figure 6.2:

1. **Clarification**: the user is shown a picture of the avatar and is explained who it is and how it behaves. The experimenter checks that the user understood correctly by asking questions, e.g. "How would you describe this character?". The experimenter then orally validates or corrects the answers of the user.

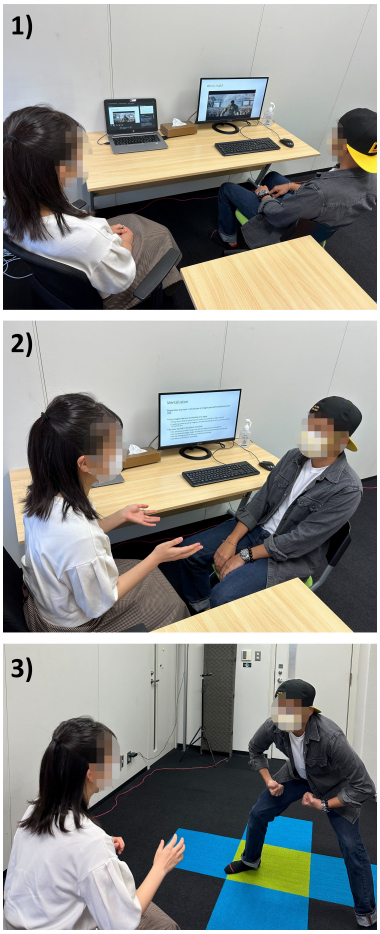


Figure 6.2: Mental preparation process. 1) *Clarification* After seeing a picture of their avatars, the participant watches a video to better understand who it represents and how it behaves. 2) *Mindfulness* The participant imagines becoming their avatars by listening to what the experimenter describes while closing their eyes. 3) *Acting* The participant imitates the avatar's behavior while following the instructions of the experimenter.

2. **Mindfulness**: the user actively imagines being the character that their avatar represents. To help them, the experimenter provides instructions on what to imagine, both visually and emotionally, e.g."Think of something that made you angry or stressed recently", "Imagine that your body is growing muscly", etc.

3. **Acting**: the user imitates the character that their avatars represent. They also receive instructions to help them perform. For example, "Clench your fists and breath deeply". After going through this step, virtual avatarization is immediately applied.

We believe this protocol may be interesting to test in both AR and VR. Indeed, the problems of avatar acceptance are similar in these two mediums and both can potentially benefit from such a preparation. To investigate its effects on the SoE, we ran a user study in VR comparing the user's SoE with and without its application. The choice to use VR rather than AR was motivated by the necessity to build a state-of-the-art embodiment system and by the larger amount of research to compare our results.

6.2 User study

The goal of this experiment was to start testing the approach described in Section 6.1.2. More generally, we wanted to check if mentally preparing users to embody their avatars could lead to a stronger SoE and/or help reduce inter-subject variability. We additionally aimed to better understand the mechanisms that lead users to accept their avatars and raise new questions about the conditions that favor strong BOIs.

6.2.1 Apparatus

System. We used an HTC Vive Pro connected to a PC equipped with an NVIDIA GeForce RTX 2080 Ti graphics card, an Intel Core i9-9900K processor, and 32 GB of RAM. The HTC lighthouses were placed at opposite corners of a rectangle-shaped room of about 9 square meters. The participants were equipped with the VR headset, two hand controllers, and two Vive trackers attached to their feet to perform a 5-point body tracking. The experimental software was built on Unity 2020.1.4f1 using SteamVR and the plugin FinalIK v2.1 developed by Pärtel Lang. A C# script was used to record the hand controller's motion. To display questionnaire material inside the VR environment and record answers, we used the VR Questionnaire Toolkit developed by Feick et al. (2020).

Avatar. For our experiment, we needed an avatar model representing a character that would not be hard to imitate for people uninitiated to acting. A character that is already familiar to the user and that displays a simple behavior would therefore be a plus. We chose to use the Hulk from the Marvel franchise who has the advantage of being

a very stereotypical character, marked by his anger and brutality. He is worldwide known and seems relatively easy to imitate compared to other characters whose body languages are more subtle. It was therefore an interesting choice for our experiment.

The model we used was a free asset designed with Maya by Muhammad Awwad (Figure 6.3, right). It was rigged for animation and included detailed body textures. We created animation clips to put the user's fingers into a fist upon pressing the hand controllers' triggers. The rest of the body was entirely animated with body tracking and IK. To do so, the avatar was calibrated in a two-step process. First, the user was asked to stand up straight while wearing the VR equipment. The avatar was then scaled down to match the user's height, and retrieved by using the headset's tracked position. Next, we used the VRIK Calibration script of the FinalIK plugin to calibrate the avatar's arms and legs lengths, as well as the user camera position. Manual adjustments were quickly made by moving the transforms of the target points of the head, hands, and feet if they were not perfectly aligned with the user. Lastly, the resulting avatar was scaled back to Hulk's model size (2.5 meters tall).

Environment. The virtual environment was a realistic industrial area created with free assets (Figure 6.3, left). Users were told they could walk, but only in a restricted area delimited by a green circle on the floor. This was to avoid collision with real surfaces. The environment included life-size objects such as soda cans to provide a visual reference for users to understand the size of their bodies and the environment's content. A mirror was placed in front of them to let them observe their entire bodies. Two conveyor belts that would carry breakable stones for the experiment's main task were placed on their right (see Figure 6.6). We included ambiance noise (low sound of construction work), as well as sound and vibration feedback provided by the hand controllers when touching the stones. This feedback varied with the strength of the touch applied.

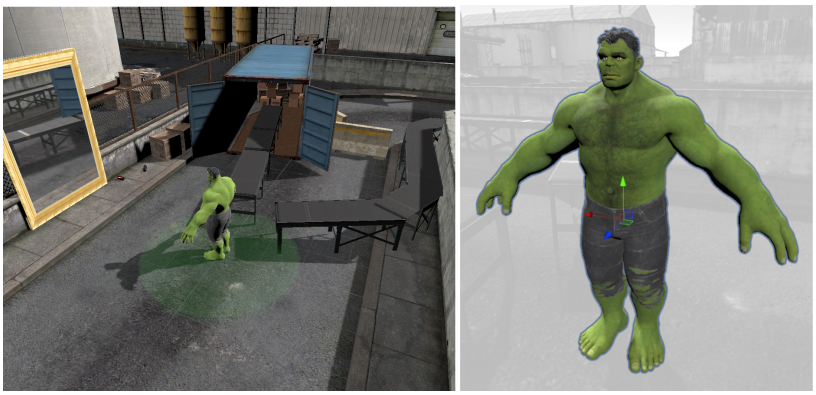


Figure 6.3: *Left* Global view of the VR environment. It includes a mirror, two conveyor belts, and diverse objects recreating an industrial scenery. The green zone on the floor was the area where users were allowed to walk. *Right* Hulk model used in the experiment. Available at: https://www.3dart.it/en/free-hulk-rig-3d-model/

6.2.2 *Methodology*

We compared two conditions in a within-subjects design: one where the users performed an arbitrary task before the embodiment, and one where they went through the mental preparation procedure. Once in the VR environment, both conditions were exactly the same. We ran these two conditions in a within-subject design as recommended by Richard et al. (2022). Since it was likely that order effects would occur, we split the conditions into two sessions separated by one week. The conditions were the following:

Figure 6.4: Participant filling a Sudoku sheet during condition NO_PREP. If this was the second session, this task was meant to prevent them from thinking about the mental preparation they had been through the previous time.

- **Condition "NO_PREP" (no preparation)**: control condition where the participant sat down and filled Sudokus on paper for 12 minutes (see Figure 6.4). They were told that their Sudoku performance was not assessed and that they would embody an avatar representing Hulk in VR after that (without further details). Once the Sudoku phase was done, the participants put on the VR equipment and were asked to close their eyes while the experimenter launched the software and calibrated their avatars. Finally, they went through the VR phase.
- **Condition "PREP"**: participants sat down and were explicitly told that their goal was to imagine being Hulk as strongly as possible. They were briefed about the preparation process and about the Hulk avatar they would embody in VR afterward. Then, they were equipped with the feet trackers and went through the three-step process described in Section 6.2 for about 12 minutes. The detail of the instructions is in Appendix B. Finally, they were equipped with the rest of the VR equipment and were asked to close their eyes while the experimenter launched the software and calibrated their avatars. The same VR phase as condition NO_PREP then began.

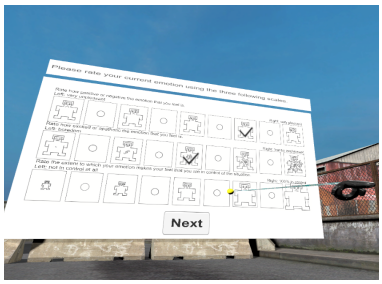


Figure 6.5: VR questionnaire interface that was used to measure the user's emotional state, the SoE, and cybersickness symptoms.

During the **VR phase**, the participants were first given a short moment to get familiar with the environment. They could see their Hulk body in the 1PP (Figure 6.6, middle). The experimenter asked them to face the virtual mirror and reproduce the gestures displayed in a video appearing next to it (Figure 6.6, left). The gestures were simple arm and leg movements (raising arms, lifting legs, etc.) repeated for 2 minutes to provide VM stimulation. After that, the experimenter asked the participants to face the conveyor belts and to try breaking the stone carried by them (Figure 6.6, right). A total of 30 stones were carried to the user for 1.5 minutes. If the participant punched strong enough[3], the stone would break into pieces with a shattering sound and a large controller vibration.

Lastly, the participants were asked to turn around and face the questionnaire interface brought up by the experimenter. They could select answers using hand rays activated for this purpose. Once done, the experimenter stopped the experimental software and unequipped the user before inviting them to sit again and fill out more questionnaires.

[3] This threshold was experimentally chosen while prototyping the system.

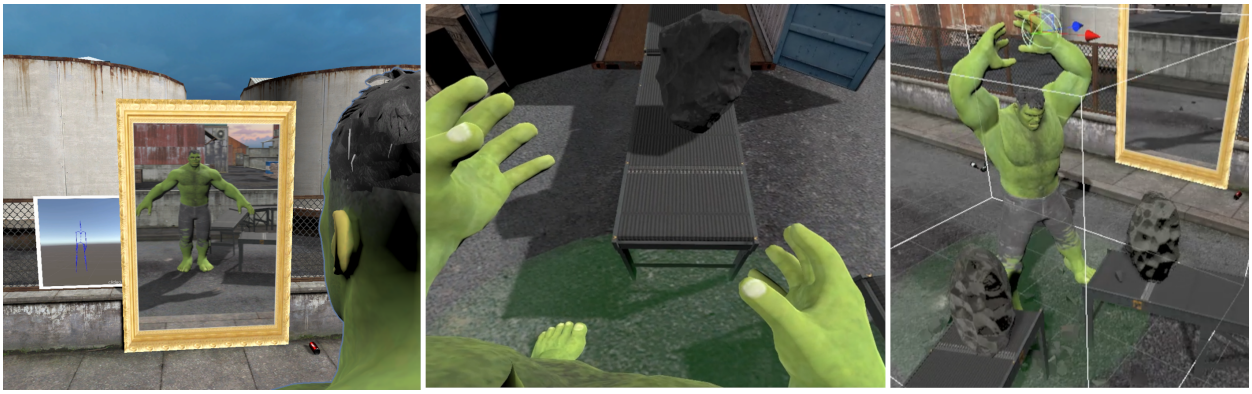


Figure 6.6: *Left* VM stimulation while facing the virtual mirror. *Middle* 1PP view of the participant. *Right* Scene view during the main task where users had to break stones. The green circle on the floor shows the area that users were instructed to stay in.

6.2.3 *Measures*

We collected demographic data to characterize the sample population: age, gender, previous experience with VR technologies, videogame play time per week, previous participation in acting activities, and level of familiarity with the character of Hulk. During the experiment, we asked the participants to answer several questionnaires detailed in Appendix B:

- **Sense of embodiment**: assessed with the revised 16-item questionnaire of Peck and Gonzalez-Franco (2021). The items were adapted to our experiment's context and applied during the VR embodiment after the punching task.
- **Emotional state**: measured with the Self-Assessment Manikin (SAM) scale by Bradley et al. (1994). This scale evaluates the valence, arousal, and dominance dimensions of a person's affective state. We applied it upon the participant's arrival, during the VR embodiment (after the punching task), and after leaving the virtual environment.
- **Simulator sickness**: assessed with the 16-item questionnaire of (Kennedy et al., 1993) to control cybersickness as a potential confounding variable.
- **Subjective feedback**: obtained through a 7-item custom questionnaire on a 7-point Likert scale. It also contained an optional comment box for participants to write open-ended feedback and was applied at the end of each session, outside of the virtual environment.

The emotional state was measured to control if the participants could successfully imagine the emotions described by the experimenter during the preparation phase. Given they would embody Hulk, they were asked to imagine being very angry and stressed as this is one of the main traits of this character. We expected the SAM applied during the embodiment would show lower valence, higher arousal, and higher dominance scores than the SAM applied upon

the participant's arrival. We also expected the SAM applied after the embodiment to show higher scores in condition PREP. In addition to these questionnaires, we recorded the following measures:

- **Punch speed**: speed of the avatar's hand upon entering in contact with a stone. Recorded for each stone touched.
- **Hand motion**: distance traveled by each hand controller during the punching task.
- **Behavior checklist**: sheet completed by the experimenter when observing the participants during the VR phase. It was meant to keep track of the user's behavior. The list included the following checkboxes: arches back, bends knees, takes a fighting posture, stomps, kicks debris, breaths heavily, and clenches fists.

The punch speed was recorded to verify if the mental preparation led the users to punch stronger. Similarly, the hand motion was recorded to see if they would make more ample movements or be more agitated in condition PREP. As the users were not specifically instructed to behave in any particular way once in the VR environment, these two measures served to indicate whether the preparation process had lasting effects on the user. The checklist provided additional information about the type of movements that they made on their initiative to embody their avatars.

6.2.4 *Participants*

We recruited 13 participants from age 25 to 35 ($\overline{m} = 29.5, SD = 4.1$). Only male participants were selected to match the avatar's gender. Among them, 53.8% had used a VR headset once or twice before, 23.1% used them a couple of times per year, 15.4% never, and 7.7% had a frequent usage history. Regarding their participation in acting activities (theater, role-playing, etc.), 61.5% had never done any before, 30.8% did some once or twice, and 7.7% practice them a couple of times per year but not more than monthly. All participants knew about Hulk: 61.5% indicated knowing him well from movies or comics, 30.8% had watched at least one movie featuring him, and 7.7% had never watched any content featuring Hulk but to know about him.

The participants were paid 1000 JPY in cash after the second session. All were unaware of the purpose of the experiments. Lastly, this experiment was approved by a local ethical committee.

6.2.5 *Results*

Embodiment. The answers to the SoE questionnaire were used to compute the following five scores: *Appearance, Response, Ownership, Multisensory,* and *Global Embodiment*. These scores were obtained through the procedure described by the questionnaire's authors (Peck and Gonzalez-Franco, 2021). Table 6.1 and Figure 6.7 summarize the results of our descriptive analysis of the SoE.

Our conditions did not show large differences in SoE scores. We checked the normality of the scores with a Shapiro-Wilk test and confirmed the data met the assumptions for a parametric test. Therefore, we applied repeated-measures T-tests on each score. We could not find any significant difference between our two conditions.

Table 6.1: Descriptive statistics of the embodiment scores. We report the median and the first (Q1) and third (Q3) quartiles for each item in the following format: Median[Q1,Q3].

condition	Appearance	Response	Ownership	Multisensory	Global SoE
NO_PREP	4.38[3.38, 5.19]	4.58[3.67, 5.58]	5.08[3.75, 5.71]	5[4.21, 5.5]	4.87[3.75, 5.52]
PREP	4.5[4.09, 5.38]	4.83[4.33, 5.54]	5.58[4.79, 6.08]	5.42[4.83, 6]	5.03[4.34, 5.73]

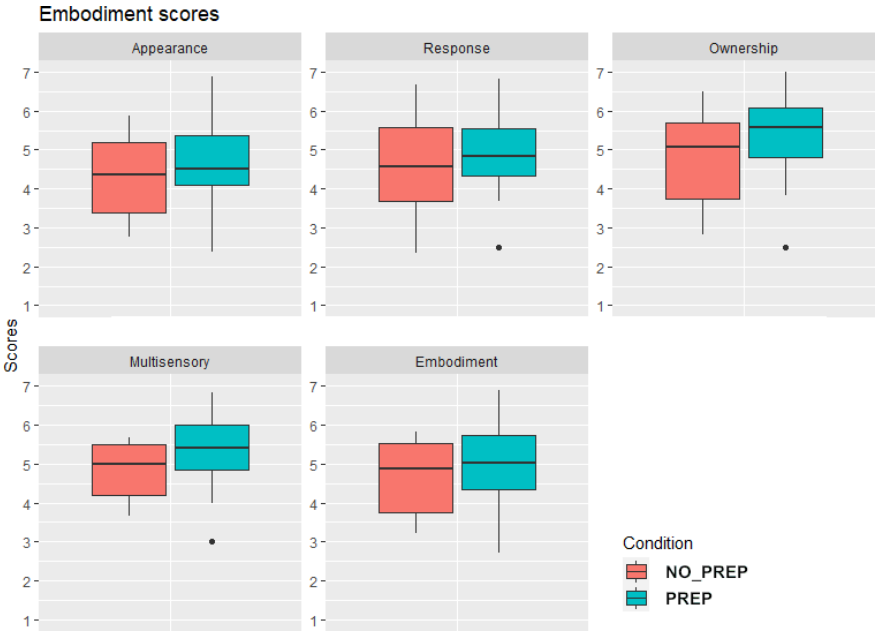


Figure 6.7: Boxplots representing the embodiment scores for each condition.

Post-experiment questionnaire. Table 6.2 and Figure 6.8 outline the results of our descriptive analysis of the post-experiment questionnaire results. This feedback shows more visible differences between our conditions, notably regarding P2, P3, P5, and P6.

Given the ordinal character of the data, we applied Wilcoxon rank sum tests on each item. The results show that participants gave significantly higher ratings to item P2 in condition PREP ($p = 0.0237$*) with a moderate effect size (r = 0.468). This indicates that the mental preparation led them to feel stronger or more powerful while embodying the avatar compared to when there was no preparation. Our analysis also revealed significantly higher ratings for item P6 in condition PREP ($p = 0.0346$*) with a moderate effect size (r = 0.437). This suggests that users felt more immersed in their avatars when they had received the mental preparation compared to when they had not.

Table 6.2: Post-experiment questionnaire. We report the median and the first and third quartiles for each item.

ID	Questions	NO_PREP	PREP
P1	"I felt my body was taller while in the avatar's body."	5[4, 6]	5.5[5, 6.25]
P2	"I felt stronger or more powerful while embodying the avatar."	5[4, 6]	6[5.75, 7]
P3	"I felt angry or wild while embodying the avatar."	3.5[2, 4.25]	6[3.5, 6.25]
P4	"It seemed to me as if I was just in a costume while being my usual self."	5.5[4.5, 6]	4.5[3, 6]
P5	"I felt engaged when punching the stones."	6[4.75, 6.25]	6.5[6, 7]
P6	"I momentarily forgot I was in an experiment."	5[2.75, 6]	6[5.75, 7]
P7	"I was conscious of the physical environment around me."	6[4.75, 6]	4.5[4, 6]

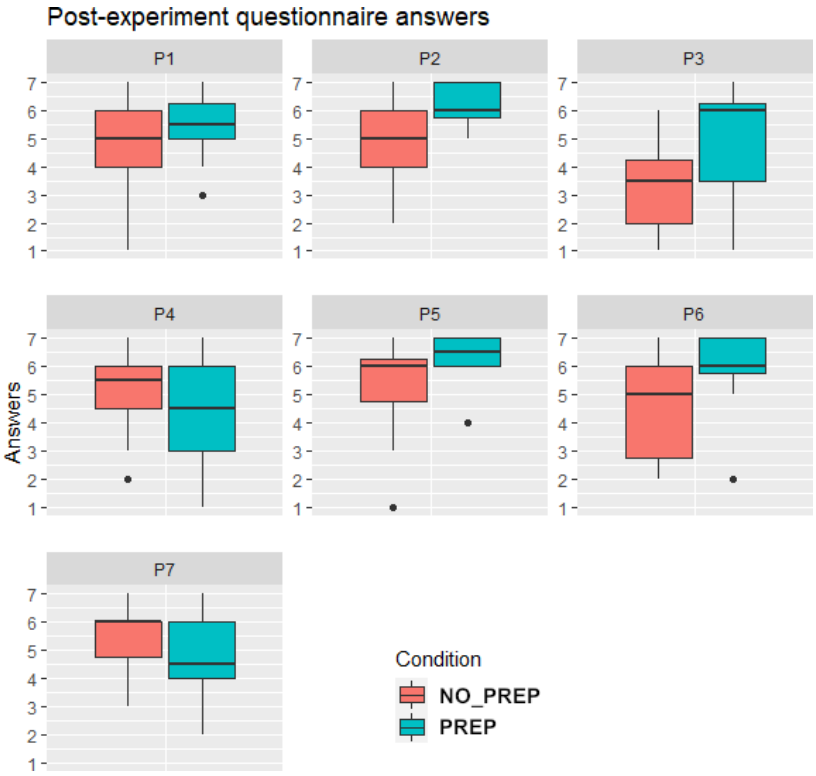


Figure 6.8: Boxplots representing the answers to each item of the post-experiment questionnaire.

Other measures. The hand motion, punch speed, behavior checklist, emotional state, simulation sickness, and the relations between them were not analyzed at the time this manuscript was submitted.

6.2.6 *Conclusion*

As the analysis of this experiment is still in progress, it is impossible to provide a global discussion of its outcomes. In light of these preliminary results, however, it seems that we will not be able to confirm our hypothesis that our preparation protocol would strengthen the SoE. Despite the one-week separation between the sessions, we suspect that order effects might be linked to this, as well as the small number of participants. Subjective feedback nevertheless suggests that our approach may have improved the user experience of their Hulk avatars. Other effects remain to be investigated and will be the object of continuing research after the submission of this thesis.

6.3 *Graphical transitions for body avatarization*

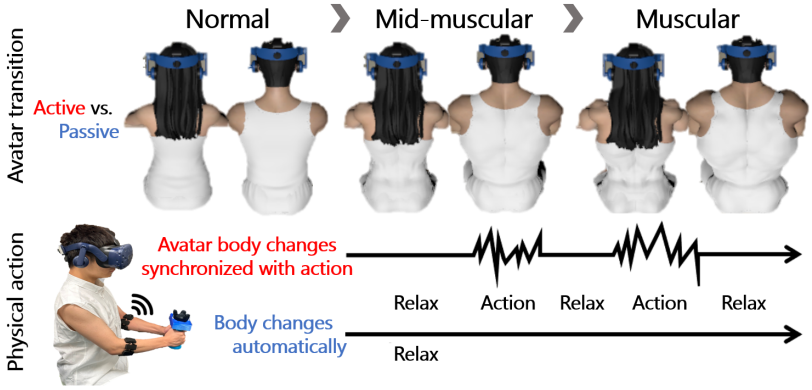


Figure 6.9: Illustration of the transitions applied, synchronized with the user's physical actions detected based on surface electromyography magnitude.

Beyond the work presented above, we also participated in research conducted by Riku Otono at the Care Lab. His idea was to apply **visual effects** at different stages of the user's avatarization to make it feel smoother. In addition, he proposed to first embody an **intermediary avatar** that looks similar to the user before switching to the final avatar. As it is not clear which types of transitions should be used, we conducted a user study comparing two of them: active and passive transitions. In both cases, visual effects are applied to let users see their avatar's body gradually change. However, during **active transitions**, the user is in control of the progress of this change through their physical action whereas during **passive transitions**, this change is applied automatically. On top of comparing these two transitions, our experiment included two other conditions in which a black screen was displayed during each type of transition to verify the impact of seeing visual effects.

The user study was conducted using an AR system letting users embody a generic avatar that transforms into a muscular body avatar in the 3PP (see Figure 6.9). We measured the user's SoE and Surface Electromyography (sEMG) magnitude to evaluate their physical performance and identify potential Proteus Effects. We hypothesized that the SoE would be stronger in the presence of avatar transitions. We supposed this would then lead users to apply more force during a physical task (see Figure 6.10) and therefore result in higher sEMG magnitudes.

Our results suggest that graphical transitions do improve the user's SoE and increase the Proteus Effects of muscular avatars. In particular, we found that the Proteus Effect was sustained longer in the presence of active transitions. We concluded that visual effects synchronized with physical action are efficient to let users feel agency and benefit from a longer Proteus Effect during embodiment. This phenomenon could be used in AR training for sports by employing a muscular avatar or for rehabilitation scenarios. We encourage future work to pursue this line of research with other types of avatars to verify the extension of these results.

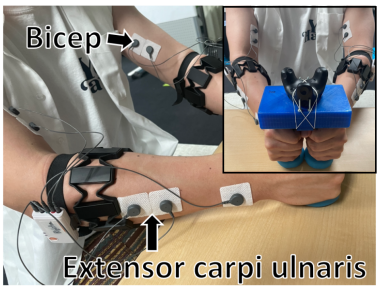


Figure 6.10: Physical task performed by the users to evaluate Proteus Effects: users had to grip these dumbbells while sEMG electrodes recorded their muscle activity.

If you can see the light at the end of the tunnel,
you are looking the wrong way.

Barry Commoner

7
Discussion and Perspectives

Throughout this Ph.D., we investigated the embodiment of virtual avatars in AR. Our main objective was to better understand how users perceive these avatars, how this perception can benefit them in daily life scenarios, and how to strengthen embodiment illusions. To pursue this goal, we accomplished the following:

- We formalized a taxonomy of "avatarization" by introducing a body avatarization continuum.
- We studied the impact of mixing virtual and real content on body ownership illusions in optical see-through AR.
- We proposed and implemented the concept of using avatars in the third-person perspective to explore the real world around us.
- We evaluated how the manipulation of such avatars can enhance our perception of physical affordances.
- We designed and tested strategies to smooth the experience of body avatarization.

Before discussing of perspectives brought by this thesis, the following section will briefly summarize how we achieved these different contributions.

7.1 Summary of contributions

Our research started with theoretical work to better understand and describe embodiment experiences: in Chapter 1, we introduced a **body avatarization continuum** offering a way to categorize systems according to the extent of the user's virtual representation. Using this continuum, we then reviewed the AR literature on avatar embodiment in Chapter 2. This allowed clarifying the technical aspects and challenges linked to body avatarization in AR, the methods available to induce and measure the SoE, and the existing similarities between AR and VR embodiment. On the other hand, we identified a lack of research on how the nature of AR impacts the user's perception of their avatars.

In Chapter 3, we delved into this topic through a **user study investigating whether the mixed content of AR environments could affect BOIs**. Our experiment compared the user's SoE towards virtual robot hands embodied in the presence of different quantities of virtual and real objects. We built an OST display allowing participants to see these hands animated by their movements, overlaid on top of their own hands. Our results show that the participants felt stronger ownership of the virtual hands when controlled in the presence of virtual and real objects mixed together, compared to when the virtual hands were the only virtual content visible. This suggests that the environment's content should be taken into account when designing embodiment experiences. However, we could not conclude with confidence on the origins of the observed differences and their extent to larger populations. This experiment should nevertheless encourage the community to further investigate the idea that the avatar itself is not the only moderator of the SoE, but that the context and medium of the embodiment may also play a role.

Next, we researched how body avatarization and AR can be combined together for concrete use cases. Using cognitive theories on affordance perception, we proposed **the concept of sending one's avatar to act remotely and improve our understanding of the physical world**, as well as of our action possibilities within it. Chapter 4 details the proof-of-concept system we implemented to illustrate this idea. It allows manipulating an avatar of matching body shape and size with three complementary techniques, namely Physical Control with a gamepad, Puppeteering with mid-air interactions, and markerless Body Tracking with an RGB-D camera. We described scenarios where each of these techniques can be relevant and discussed their individual strengths and weaknesses.

Chapter 5 continued this research by **evaluating the effects of 3PP avatar control on real-world affordance perception**. We ran three studies where participants controlled their self-avatars to test actions in the physical environment in front of them. The first experiment showed that manipulating a virtual version of oneself can effectively improve one's accuracy and confidence to estimate action capabilities. The second experiment investigated how the memory of the avatar's manipulation can be exploited when reiterating the same judgments once the avatar is no longer in sight. We could not find evidence that having used the avatar impacted the user's performance as we found no significant difference between the conditions that included prior usage of the avatar and the ones that did not. The last experiment continued our inquiry into the resilient effects of avatars by asking users to estimate a doorway's passability. It additionally compared their accuracy when facing a real door to when facing a virtual one. We could not find significant effects of the avatar, nor of the door's virtuality on the user's performance. Regardless, from Experiment 1, we retain that the experience acquired through the avatar can still be used to make better-informed decisions and to better grasp the dimensions of things.

Lastly, we investigated **strategies to strengthen BOIs and reduce variability** caused by individual user characteristics. Chapter 6 describes our collaborative research with the CARE lab on how to exploit the different phases of embodiment experiences, and in particular, the time spent before the embodiment. We presented a VR pilot study evaluating how the user's mental conditioning can help to create stronger BOIs. Participants went through mindfulness and acting exercise that we had designed to help them imagine becoming their avatars before embodying them. We evaluated the effects of this preparation on the SoE toward an avatar representing the Hulk. The results of this experiment could not be fully presented as their analysis was still in progress when submitting this thesis. However, a preliminary look seems to indicate that this mental preparation might have led users to feel more powerful while embodying Hulk and more immersed in the VR environment. Lastly, we participated in the research of Riku Otono on transition effects to smooth the experience of body avatarization. The results of this work indicate that applying visual effects and using an intermediary avatar that looks more similar to the user can enhance the sense of agency and extend Proteus Effects. We concluded that such strategies can be effective to reinforce BOIs and deserve to be studied in more depth in future work.

7.2 Short-term perspectives

Much remains to be addressed before we fully understand how avatar embodiment is experienced in AR. In this section, we discuss the open questions and perspectives highlighted by the contributions of this thesis.

7.2.1 On the impact of AR content on the sense of embodiment

The user study of Chapter 3 suggested that environmental factors can affect BOIs in AR. However, our secondary analysis did not allow us to definitely confirm that the mix of real and virtual content actually affects the SoE. We concluded that this was probably due to our relatively small number of participants, but there is a possibility that the differences we had initially observed were not linked to our conditions. This would imply that the SoE was built independently of the environment's content in our experiment. Although it is impossible to assert that mixing contents really makes no difference, two papers are consistent with this interpretation.

The first is that of Škola et al. (2016) who compared the traditional BOI and the VHI in both VR and AR. Their findings show that AR and VR produce similar illusions, albeit weaker than with a real rubber hand. These results support the idea that AR and VR can be used interchangeably to create BOIs. The second paper is that of Wolf, Döllinger, et al. (2020) who compared full BOIs in VR and AR with

a VST HMD[1]. Unlike what the authors were expecting, their participants did not report substantially different senses of ownership and presence in AR and VR. They speculate that the immersion provided by their systems was maybe too similar, i.e. the VST video resolution and VR graphics might have been too comparable for users to experience a difference.

While this explanation makes sense and supports their initial hypothesis that the SoE should differ in AR, this lack of difference can be equally assumed as a sign that the BOIs evolved independently of the displayed context. Establishing that there is no difference between BOIs in AR and VR, however, remains impossible. Using OST instead of VST would maybe have changed the results, as might having recruited more participants. Until a larger set of experiments is run, it cannot be assumed that the display media and context have no play at all. Conversely, if future work confirms that environmental context is impactful as suggested by our experiment, questions will remain on how exactly it affects the SoE, in which circumstances, and how this is related to other factors. We encourage future work to push the investigations further on this subject.

[1] Subjects saw a photorealistic generic avatar through an interactive virtual mirror that offered visuomotor feedback. The physical world viewed in the VST condition was faithfully reproduced in the VR condition.

7.2.2 *On the implementation of AR embodiment systems*

Although more and more papers are implementing AR embodiment systems, research seeking applications for them is still at its dawn. We explored new functions for self-avatars in AR and proposed the concept of using them to better understand the world that surrounds us. Chapter 5 showed that building such AR systems is already possible. However, we were confronted with both technical difficulties and multiple interrogations regarding how to design it.

First, our developments made us realize that there was a significant lack of guidelines for designing AR embodiment applications. In Chapter 4, to help future work, we attempted to detail our creation process and outlined different approaches to control avatars in the 3PP based on different needs. However, the control of the avatar is only one dimension of embodiment experiences; the user perspective, extent of avatarization, abstraction, rending type, and other aspects have to be considered when creating an avatarization system. Identifying these dimensions of embodiment experiences and constructing a design space could therefore benefit the creation process of future AR embodiment applications.

Second, our developments were subject to several technical limitations. For example, while prototyping, we decided to remove the physics simulation service provided by the Hololens 2 because of its lack of accuracy, which would have caused too many discrepancies. In addition to encouraging the formalization of guidelines, we therefore suggest future work also focus on solving these limitations. Many difficulties remain to be addressed regarding the proper 3D registration and interactivity of avatars. In particular, full avatarization in the 1PP has barely been approached due to the difficulty

of displaying virtual overlays in real time. Improving the technical knowledge and providing implementation examples would facilitate the exploration of the potential of AR avatars, but would also enable the development of new studies to better understand their perception.

7.2.3 *On the perception of self-avatars in AR*

Although it was not the general tendency, the user study of Chapter 5 showed that some participants were reluctant to rely on their avatars to judge their possibilities of action. Given they were explicitly told that their avatars were customized to match their real bodies, this observation rose several questions about the trust that users put in their avatars.

Looking at self-perception literature, it seems that healthy people often perceive their body sizes inaccurately (Piryankova et al., 2014). The study of Nimcharoen et al. (2018) showed that this mismatch can also apply when presented with self-avatars in a virtual mirror as participants tend to identify themselves with slightly thinner virtual bodies than their own (90% of their width). It is possible that the participants of our experiment experienced a similar bias and that it caused discrepancies when seeing that their avatars could reach objects farther than they thought. This mismatch of expectations was expressed by many participants, but it seems most still decided to rely on their avatars during the evaluation.

In future work, it would be interesting to test up to which point the avatar — at first faithful to the user's body shape and size — can be deformed before the user realizes it. The link between the strength of the SoE and this breaking point could also be relevant to investigate: would a strong embodiment help users to accept more important body modifications? A future study could additionally assess whether changes going unnoticed by the user could influence the perception of one's capabilities. For example, does embodying a larger avatar make users underestimate their ability to enter a narrow space? If so, it could be interesting to examine using these artificial biases to correct our perception when it is faulty.

7.2.4 *On the variability of embodiment illusions*

Investigating the influence of avatar appearance in AR was out of our scope, but studying it further may help us understand some of our results. Throughout our experiments, we could observe that the avatarization we applied evoked highly varying responses from the participants. In Chapter 3, for example, several participants reported being estranged from the virtual robot hands they embodied, as if they were gloves rather than their actual hands. Similarly, although it was not the overall trend, some participants of Chapter 6 described that they felt as if they were in a costume rather than in the body of Hulk. It is difficult to understand why only certain people had this

impression and which aspects of the 3D models caused it.

In the case of Hulk, it is possible that the willingness of users to embody their avatars depended on their affinity with the character it represented. Participants who originally disliked the Hulk character might have been reluctant to accept its body as theirs. However, the dominance of such top-down factors has not yet been investigated to our knowledge, whether it be in AR or other research. Secondly, it seems plausible that an avatar appearing lifelike to one person shows significant flaws to the eyes of another, hence ruining the illusion[2]. This phenomenon was often described in previous VR experiments, showing highly variable results with similar types of avatars (Lin et al., 2016; Lugrin et al., 2015; Schwind et al., 2017).

[2] For example, a gamer who is used to playing games in ultra-high definition may see their expectations disappointed when presented with an avatar with less realistic rendering, compared to an XR user who is used to this type of graphics.

Lugrin et al. (2015) hypothesized that this variability could be due to the Uncanny Valley phenomenon, conceptualized by Mori et al. (2012). According to the theory, users feel a greater connection to anthropomorphic avatars as their degree of human resemblance rises until it reaches a point where the connection is brought to a sudden decline. This connection is reestablished to a greater extent from the point where the user considers the avatar to be realistic enough. Such judgment being highly subjective would explain why avatars produce illusions that vary largely with users.

At this point, it is unclear whether experimenters seeking strong embodiment illusions in AR should privilege certain appearances. As experimental results on this subject are contrasting, it seems crucial to continue research that will help understand their effects. We suggest that future work investigates how subjective factors (familiarity, suggestibility, personality, etc.) impact the way users react to different avatars.

7.3 *Long-term perspectives*

We are still a long way from reaching the Ultimate Display and the Avatar Dream. Obtaining technologies equalling those of the movie of Avatar seems even farther away: in this movie, the main character can control an organic body without moving and feel sensations through it. Current AR technologies allow seeing oneself inside a virtual body and controlling it within physical space, but they support only limited interactions and provide rudimentary tactile feedback. Therefore, the next big step seems to be the creation of multisensory environments that not only involve the realistic integration of avatars in the real world but also touch sensations when embodying them. The skin forms the basis of our physical interactions with the world and allows us to define a mechanical understanding of our environment. By integrating active tactile feedback, virtual content becomes tangible and truly integrated into the real world, as physical objects.

Attempts to incorporate tactile sensations into AR applications are still in their infancy. This is probably due to the slow diffusion of haptic interfaces adapted to AR that are user-friendly, lightweight, and affordable. Rosa (2016) also outlines the under-exploration of

other cues such as audio feedback in AR experiences. For instance, integrating heavy footstep sounds when embodying a weighty avatar could improve the credibility of the embodiment. More generally, visual and tactile stimulation provides users with natural interaction feedback that can greatly enhance their understanding of the AR system's dynamics.

Secondly, important progress needs to be made in Brain-computer Interface (BCI) research before users can control a whole virtual body without moving, let alone a biological one. The investigation of avatar control through BCIs in AR, therefore, offers a second long-term view for avatar improvement. Important progress in visual tracking technologies now allows applying a one-to-one mapping of user movements onto their avatars without requiring them to wear trackers. However, the agency is still limited by the number of tracked joints and accuracy remains insufficient to provide seamless animation. Additionally, AR HMDs are hardly incompatible with facial tracking, and avatars rarely represent well their user's expressions. Until BCIs reach the required maturity to provide a fine degree of control of avatars, research will have to seek alternative ways to make avatar animation more representative of their users.

It will probably take many lifetimes to get James Cameron's avatars —if we ever get there. Nevertheless, the next generation of avatars can be expected to arrive much sooner and will already allow us to do much more than today. We propose to discuss some of the perspectives that this new generation may bring in the following paragraphs.

7.3.1 *Promises of AR avatars in the medical field*

One of the sectors that could highly benefit from the development of AR avatars is medical care. With sufficient progress in BCIs, physically disabled people could become more autonomous by controlling virtual limbs to interact with connected devices (Feuchtner et al., 2017). Patients preparing for graft could also train themselves to use their prosthesis in advance by seeing it virtually attached to their bodies, as (Nishino et al., 2017) started to investigate (Figure 7.1). Aside from prosthetics, new treatment strategies could also be developed in various medical contexts: pain treatment, tremor stabilization, and stroke rehabilitation using AR avatarization are already being studied (Eckhoff et al., 2019; Kaneko et al., 2019; K. Wang et al., 2017). Lastly, all of this could be done from home as the doctors would be able to virtually visit their patients and incorporate therapies right into their environments. Not breaking the normal course of a day might lead to higher acceptability of treatments and psychotherapies, but also finds utmost interest in the assistance of mentally impaired people (e.g. autism spectrum disorders (Kellems et al., 2020)).

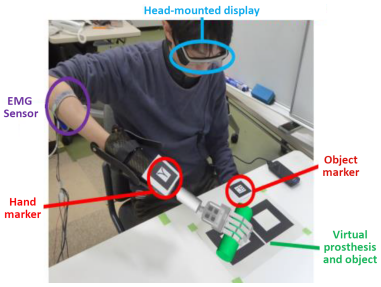


Figure 7.1: Development of a myoelectric prosthesis simulator using AR by Nishino et al. (2017).

7.3.2 *Learning from AR embodiment experiences*

Harnessing the Proteus Effect in AR could open additional doors to novel educational methods and skill enhancement. Indeed, VR studies previously showed how embodying avatars representing stereotypes of intelligence or musicality could momentarily improve the user's executive functioning, problem-solving skills, artistic and motor abilities[3]. One can easily picture the advantages of exploiting these gains within specific real-world contexts by selecting the right full avatarization. Alternatively, the Proteus Effect could be used to increase awareness and reduce bias. Research has shown that participants could experience a reduction of implicit biases after embodying avatars with different gender, race, or ethnicity in VR (Banakou, Hanumanthu, et al., 2016; Beltran et al., 2021; Peck, Seinfeld, et al., 2013). This reduction could potentially be exploited for a long time over normal life thanks to AR.

[3] Banakou, Kishore, et al. (2018) showed that embodying Albert Einstein in VR could increase executive functioning and planning abilities. Kilteni, Bergstrom, et al. (2013), on the other hand, showed that embodying a dark-skinned casually dressed avatar resulted in increased movement patterns while playing the drums than with a light-skinned formally dressed avatar.

7.3.3 *Role-playing and AR gaming*

AR avatarization has unequivocal potential in innovative storytelling and role-playing games. Through AR, users could instantly step into the shoes of a story's main character without having to immerse themselves in a virtual world. Taking such gaming experiences outside of VR allows seeing the physical world and exploiting its configuration. This brings supplementary opportunities for appealing game-plays involving interaction with real-world objects (e.g. hiding behind real walls from attackers) and people (Nilsen et al., 2004; Thomas, 2012). Although they have their share of technical difficulties, AR visualizations are not concerned with many of the challenges linked to VR gaming, including avoiding real-world obstacles (e.g. chairs), dealing with cybersickness, or navigating in large spaces with room restrictions. Currently, most AR games are handheld device applications that use screen interactions, leaving body avatarization aside (Chatzopoulos et al., 2017). Although mobile AR is an easy entry point, we believe HMDs will allow achieving the true potential of AR gaming by letting users embody game characters and customize their appearance.

7.3.4 *AR avatars as vehicles of the self*

Lastly, in a hypothetical future where connected devices are ubiquitous and smartphones are replaced by AR glasses, people could walk the street and display their avatars on their bodies for others to see. This would enable people to explore and express their identity more easily by allowing them to customize their appearance virtually at will and without limit. Looking at the advent of virtual influencers and similar entities, it seems that modern-day social media consumers and companies are already heading in this direction. Some users design social identities from scratch and materialize them as hyper-realistic avatars for which they create online

profiles or video channels. Rapidly gaining in popularity, some of these profiles gather over 3 million followers and sometimes present videos of flesh-and-blood people interacting with the featured avatar (e.g. 19-year-old LilMiquela on Figure 7.2). The decrease in the cost of tracking solutions and AR devices will eventually lead to more avatar-people live interactions. Regular users will then be able to embody their personalized self-avatars and convey the image they choose to be and to see of themselves.

Figure 7.2: "Que sexy". LilMiquela via Instagram.

7.3.5 *But is it right? Ethical questioning*

As the ideas we describe may become possible in a near future, it seems urgent to better assess perceptual and cognitive characteristics of the embodiment of avatars in AR, but also to evaluate the medical, social, and ethical risks of letting people take ownership of appearances that are not their own. Indeed, achieving the Avatar Dream might turn out harmful if it reinforces current social biases or tampers with human well-being, property, and privacy rights. Future studies should strive in making this dream a positive one, fostering self-fulfillment and encouraging face-to-face interaction instead of degrading human relations.

Secondly, there is a possibility that embodying avatars through AR allows for experiencing longer-lasting effects than what was observed in VR. This idea was previously described by Eckhoff et al. (2019) who hypothesized that realistic AR experiences could induce stronger virtual synesthesia than in virtual environments: being applied in a world known to be "real", illusions could be perceived more easily as actually *being* happening. This possibility is of high interest for many applications, but it also raises many thought-provoking questions about the long-term usage of avatars in AR. Can the Proteus Effect alter our behaviors beyond the scope of AR experiments? Would this alteration be durable? How can we control and use it? What risks arise out of it?

Understanding these matters will have immediate and tangible consequences for science, design, and production. However, for the time being, these interrogations remain unanswered and based on suppositions. As research has made it clear that virtual embodiment is useful for more than just decoding cognitive processes, we expect that future work will find great interest in investigating them. Until then, I hope that the research presented in this thesis will act as a small step towards an inspiring future for avatars and that others will carry on this path.

To all the researchers that came before me,
and graciously lent me their shoulders to stand on,
and without who this work could be very different.

Fabien Benureau

Bibliography

Adam, Hajo and Adam D Galinsky (2012). "Enclothed cognition". In: *Journal of experimental social psychology* 48.4, pp. 918–925.

Aglioti, Salvatore, Nicola Smania, Michela Manfredi, and Giovanni Berlucchi (1996). "Disownership of left hand and objects related to it in a patient with right brain damage". In: *Neuroreport* 8.1, pp. 293–296.

Akçayir, Murat and Gökçe Akçayir (2017). "Advantages and challenges associated with augmented reality for education: A systematic review of the literature". In: *Educational Research Review* 20, pp. 1–11.

Anania, Emily C (2019). "Patient Willingness to Undergo Robotic Surgery: Identification and Validation of a Predictive Model". PhD thesis. Embry-Riddle Aeronautical University.

Anderegg, Raphael, Loïc Ciccone, and Robert W Sumner (2018). "PuppetPhone: puppeteering virtual characters using a smartphone". In: *Proceedings of the 11th Annual International Conference on Motion, Interaction, and Games*, pp. 1–6.

Arend, Matthias G. and Jochen Müsseler (July 2021). "Object affordances from the perspective of an avatar". In: *Consciousness and Cognition* 92, p. 103133. ISSN: 1053-8100. DOI: 10.1016/j.concog.2021.103133. URL: https://www.sciencedirect.com/science/article/pii/S1053810021000593.

Arenz, D (2001). "Heautoskopie Doppelgängerphänomen und seltene Halluzination der eigenen Gestalt". In: *Der Nervenarzt* 72.5, pp. 376–379.

Argelaguet, Ferran, Ludovic Hoyet, Michael Trico, and Anatole Lecuyer (Mar. 2016). "The Role of Interaction in Virtual Embodiment: Effects of the Virtual Hand Representation". In: *2016 IEEE Virtual Reality and 3D User Interfaces (VR)*. Greenville, South Carolina, USA, pp. 3–10. DOI: 10.1109/VR.2016.7504682.

Armel, K. Carrie and V. S. Ramachandran (July 2003). "Projecting Sensations to External Objects: Evidence From Skin Conductance Response". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270, pp. 1499–1506. ISSN: 0962-8452, 1471-2954. DOI: 10.1098/rspb.2003.2364.

Ash, Erin (2016). "Priming or proteus effect? Examining the effects of avatar race on in-game behavior and post-play aggressive cognition and affect in video games". In: *Games and Culture* 11.4, pp. 422–440.

Azuma, Ronald T. (1997). "A Survey of Augmented Reality". In: *Presence: Teleoperators and Virtual Environments* 6, pp. 355–385. DOI: 10.1162/pres.1997.6.4.355.

Bailenson, Jeremy N and Jim Blascovich (2004). "Avatars. Encyclopedia of human-computer interaction". In: *Berkshire Publishing Group* 64, p. 68.

Bainbridge, William Sims (2004). *Berkshire Encyclopedia of Human-computer Interaction*. Vol. 1. Great Barrington, Massachusetts, USA: Berkshire Publishing Group LLC. ISBN: 978-0-9743091-2-5.

Banakou, Domna, Parasuram D. Hanumanthu, and Mel Slater (Nov. 2016). "Virtual Embodiment of White People in a Black Virtual Body Leads to a Sustained Reduction in Their Implicit Racial Bias". In: *Frontiers in Human Neuroscience* 10. ISSN: 1662-5161. DOI: 10.3389/fnhum.2016.00601.

Banakou, Domna, Sameer Kishore, and Mel Slater (June 2018). "Virtually Being Einstein Results in an Improvement in Cognitive Task Performance and a Decrease in Age Bias". In: *Frontiers in Psychology* 9, p. 917. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.00917.

Barbosa, Alcimar, Edgard Afonso Lamounier Junior, Adriano de Oliveira Andrade, and Alexandre Cardoso (Oct. 2012). "Virtual and Augmented Reality: A New Approach to Aid Users of Myoelectric Prostheses". In: ed. by Ganesh R. Naik. DOI: 10.5772/50600.

Baskar, Juhi et al. (2017). "HoloPHAM: An Augmented Reality Training System For Upper Limb Myoelectric Prosthesis Users". PhD thesis. Johns Hopkins University.

Batmaz, Anil Ufuk, Mohammad Rajabi Seraji, Johanna Kneifel, and Wolfgang Stuerzlinger (2020). "No Jitter Please: Effects of Rotational and Positional Jitter on 3D Mid-Air Interaction". In: *Proceedings of the Future Technologies Conference*. Springer, pp. 792–808.

Beier, Guido (1999). "Kontrollüberzeugungen im umgang mit technik". In: *Report Psychologie* 9, pp. 684–693.

Beltran, Kevin, Cody Rowland, Nicki Hashemi, Anh Nguyen, Lane Harrison, Sophie Engle, and Beste F Yuksel (2021). "Reducing Implicit Gender Bias Using a Virtual Workplace Environment". In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–7.

Bermano, Amit H, Markus Billeter, Daisuke Iwai, and Anselm Grundhöfer (2017). "Makeup lamps: Live augmentation of human faces via projection". In: *Computer Graphics Forum*. Vol. 36. 2. Wiley Online Library, pp. 311–323.

Bertrand, Philippe, Jérôme Guegan, Léonore Robieux, Cade Andrew McCall, and Franck Zenasni (2018). "Learning empathy through virtual reality: multiple strategies for training empathy-related abilities using body ownership illusions in embodied virtual reality". In: *Frontiers in Robotics and AI*, p. 26.

Bhargava, Ayush, Kathryn M Lucaites, Leah S Hartman, Hannah Solini, Jeffrey W Bertrand, Andrew C Robb, Christopher C Pagano, and Sabarish V Babu (2020). "Revisiting affordance per-

ception in contemporary virtual reality". In: *Virtual Reality* 24.4, pp. 713–724.

Böffel, Christian and Jochen Müsseler (2018). "Perceived Ownership of Avatars Influences Visual Perspective Taking". In: *Frontiers in Psychology* 9, p. 743. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.00743. URL: https://www.frontiersin.org/article/10.3389/fpsyg.2018.00743.

Bommel, JG van (2017). "Presence and Embodiment in Augmented Reality". MA thesis. Utrecht, Pays-Bas.

Botvinick, Matthew and Jonathan Cohen (Feb. 1998). "Rubber Hands 'Feel' Touch That Eyes See". In: *Nature* 391, pp. 756–756. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/35784.

Bown, Johnathan, Elisa White, and Akshya Boopalan (2017). "Looking for the ultimate display: A brief history of virtual reality". In: *Boundaries of self and reality online*. Elsevier, pp. 239–259.

Bradley, Margaret M and Peter J Lang (1994). "Measuring emotion: the self-assessment manikin and the semantic differential". In: *Journal of behavior therapy and experimental psychiatry* 25.1, pp. 49–59.

Braun, Niclas, Jeremy D. Thorne, Helmut Hildebrandt, and Stefan Debener (Nov. 2014). "Interplay of Agency and Ownership: The Intentional Binding and Rubber Hand Illusion Paradigm Combined". In: *PLOS ONE* 9.11. Publisher: Public Library of Science, e111967. ISSN: 1932-6203. DOI: 10/gh3ss2.

Cao, Zhe, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh (2018). "OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields". In: *arXiv preprint arXiv:1812.08008*.

Cardeña, Etzel and Devin B Terhune (2014). "Hypnotizability, personality traits, and the propensity to experience alterations of consciousness." In: *Psychology of Consciousness: Theory, Research, and Practice* 1.3, p. 292.

Chatzopoulos, Dimitris, Carlos Bermejo, Zhanpeng Huang, and Pan Hui (2017). "Mobile augmented reality survey: From where we are to where we go". In: *Ieee Access* 5, pp. 6917–6950.

Costantini, Marcello and Patrick Haggard (June 2007). "The rubber hand illusion: Sensitivity and reference frame for body ownership". In: *Consciousness and Cognition* 16.2, pp. 229–240. ISSN: 1053-8100. DOI: 10/cqnrb7.

Creem-Regehr, Sarah H., Kyle T. Gagnon, Michael N. Geuss, and Jeanine K. Stefanucci (2013). "Relating spatial perspective taking to the perception of other's affordances: providing a foundation for predicting the future behavior of others". In: *Frontiers in Human Neuroscience* 7, p. 596. ISSN: 1662-5161. DOI: 10.3389/fnhum.2013.00596. URL: https://www.frontiersin.org/articles/10.3389/fnhum.2013.00596/full.

Cruz-Neira, Carolina, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart (June 1992). "The CAVE: Audio Visual Experience Automatic Virtual Environment". In: *Commun. ACM*

35.6, pp. 64–72. ISSN: 0001-0782. DOI: 10.1145/129888.129892. URL: https://doi.org/10.1145/129888.129892.

Dacko, Scott G. (2017). "Enabling smart retail settings via mobile augmented reality shopping apps". In: *Technological Forecasting and Social Change* 124, pp. 243–256. ISSN: 0040-1625. DOI: 10/gh7jmb. URL: https://www.sciencedirect.com/science/article/pii/S0040162516304243.

Daiber, Florian, Felix Kosmalla, and Antonio Krüger (2013). "BouldAR: Using Augmented Reality to Support Collaborative Boulder Training". In: *CHI '13 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '13. Paris, France: Association for Computing Machinery, pp. 949–954. ISBN: 9781450319522. DOI: 10.1145/2468356.2468526. URL: https://doi.org/10.1145/2468356.2468526.

De Vignemont, Frédérique (2011). "Embodiment, ownership and disownership". In: *Consciousness and cognition* 20.1, pp. 82–93.

Debarba, Henrique G, Eray Molla, Bruno Herbelin, and Ronan Boulic (2015). "Characterizing embodied interaction in first and third person perspective viewpoints". In: *2015 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, pp. 67–72.

Dede, Chris (Jan. 2009). "Immersive Interfaces for Engagement and Learning". In: *Science* 323.5910, pp. 66–69. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1167311. URL: https://www.science.org/doi/10.1126/science.1167311.

Dewez, Diane, Rebecca Fribourg, Ferran Argelaguet, Ludovic Hoyet, Daniel Mestre, Mel Slater, and Anatole Lécuyer (2019). "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)*. IEEE, pp. 123–134.

Du, Chao, Yen-Lin Chen, Mao Ye, and Liu Ren (2016). "Edge Snapping-Based Depth Enhancement for Dynamic Occlusion Handling in Augmented Reality". In: *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 54–62. DOI: 10.1109/ISMAR.2016.17.

Durrant-Whyte, Hugh and Tim Bailey (2006). "Simultaneous localization and mapping: part I". In: *IEEE robotics & automation magazine* 13.2, pp. 99–110.

Eckhoff, Daniel, Alvaro Cassinelli, and Christian Sandor (2019). "Exploring Perceptual and Cognitive Effects of Extreme Augmented Reality Experiences". In: Beijing, China, p. 4.

Ehrsson, H.Henrik (Jan. 2012). "The Concept of Body Ownership And Its Relation to Multisensory Integration". In: *The New Handbook of Multisensory Processes*, pp. 775–792.

Eisert, Peter, Jurgen Rurainsky, and Philipp Fechteler (2007). "Virtual mirror: Real-time tracking of shoes in Augmented Reality environments". In: *2007 IEEE International Conference on Image Processing*. Vol. 2. IEEE, pp. II–557.

Everingham, M. R., B. T. Thomas, and T Troscianko (Jan. 1998). "Head-Mounted Mobility Aid for Low Vision Using Scene Classi-

fication Techniques". In: *International Journal of Virtual Reality* 3.4, pp. 1–10. ISSN: 1081-1451. DOI: 10.20870/IJVR.1998.3.4.2629. URL: https://ijvr.eu/article/view/2629.

Farrer, C, M Bouchereau, M Jeannerod, and N Franck (2008). "Effect of distorted visual feedback on the sense of agency". In: *Behavioural neurology* 19.1, 2, pp. 53–57.

Feick, Martin, Niko Kleer, Anthony Tang, and Antonio Krüger (2020). "The Virtual Reality Questionnaire Toolkit". In: UIST '20 Adjunct. Virtual Event, USA: Association for Computing Machinery. ISBN: 9781450375153. DOI: 10.1145/3379350.3416188. URL: https://doi.org/10.1145/3379350.3416188.

Fernandez-Baena, Adso, L Salle, and David Miralles (2014). "Avatar: Tangible interaction and augmented reality in character animation". In: *on Interacting with Smart Objects*, p. 28.

Feuchtner, Tiare and Jörg Müller (2017). "Extending the Body for Interaction with Reality". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver, Colorado, USA: ACM Press, pp. 5145–5157. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025689.

Fischer, Edith and Richard F Haines (1980). *Cognitive issues in head-up displays*. Tech. rep.

Fox, Jesse, Jeremy N Bailenson, and Liz Tricase (2013). "The embodiment of sexualized virtual selves: The Proteus effect and experiences of self-objectification via avatars". In: *Computers in Human Behavior* 29.3, pp. 930–938.

Fribourg, Rebecca, Ferran Argelaguet, Anatole Lécuyer, and Ludovic Hoyet (2020). "Avatar and Sense of Embodiment: Studying the Relative Preference Between Appearance, Control and Point of View". In: *IEEE Transactions on Visualization and Computer Graphics* 26.5, pp. 2062–2072.

Genay, Adélaïde, Anatole Lécuyer, and Martin Hachet (June 2020). "Incarner un Avatar en Réalité Augmentée : Revue de la Littérature". In: *Workshop sur les Affects, Compagnons artificiels et Interactions (WACAI)*. CNRS, Université Toulouse Jean Jaurès, Université de Bordeaux. Saint Pierre d'Oléron, France. URL: https://hal.inria.fr/hal-02933477.

– (2021a). "Being an Avatar "for Real": a Survey on Virtual Embodiment in Augmented Reality". In: *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1. DOI: 10.1109/TVCG.2021.3099290.

– (2021b). "Virtual, Real or Mixed: How Surrounding Objects Influence the Sense of Embodiment in Optical See-Through Experiences?" In: *Frontiers in Virtual Reality* 2. DOI: 10.3389/frvir.2021.679902.

– (Oct. 2022). "What Can I Do There? Controlling AR Self-Avatars to Better Perceive Affordances of the Real World". In: *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. Singapour. DOI: 10.1109/ISMAR55827.2022.00061.

Gervais, Renaud, Jérémy Frey, Alexis Gay, Fabien Lotte, and Martin Hachet (2016). "Tobe: Tangible out-of-body experience". In: *Pro-*

ceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction, pp. 227–235.

Gibson, James J (2014). *The ecological approach to visual perception: classic edition*. New York: Psychology Press.

Gilbers, Carolien (2017). "The Sense of Embodiment in Augmented Reality: A Third Hand Illusion". MA thesis. Utrecht, Pays-Bas.

Gonzalez-Franco, Mar and Tabitha C. Peck (June 2018). "Avatar Embodiment. Towards a Standardized Questionnaire". In: *Frontiers in Robotics and AI* 5, p. 74. ISSN: 2296-9144. DOI: 10.3389/frobt.2018.00074.

Gorisse, Geoffrey, Olivier Christmann, Etienne Armand Amato, and Simon Richir (July 2017). "First- and Third-Person Perspectives in Immersive Virtual Environments: Presence and Performance Analysis of Embodied Users". In: *Frontiers in Robotics and AI* 4, p. 33. ISSN: 2296-9144. DOI: 10.3389/frobt.2017.00033.

Griffin, Nathan Navarro and Eelke Folmer (2019). "Out-of-body locomotion: Vectionless navigation with a continuous avatar representation". In: *25th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–8.

Guegan, Jérôme, Stéphanie Buisine, Fabrice Mantelet, Nicolas Maranzana, and Frédéric Segonds (2016). "Avatar-mediated creativity: When embodying inventors makes engineers more creative". In: *Computers in Human Behavior* 61, pp. 165–175.

Haans, Antal, Wijnand A. IJsselsteijn, and Yvonne A. W. de Kort (Dec. 2008). "The effect of similarities in skin texture and hand shape on perceived ownership of a fake limb". In: *Body Image* 5.4, pp. 389–394. ISSN: 1740-1445. DOI: 10/bnc56r.

Hachet, Martin, Benoit Bossavit, Aurélie Cohé, and Jean-Baptiste de la Rivière (2011). "Toucheo: multitouch and stereo combined in a seamless workspace". In: *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pp. 587–592.

Harrell, D Fox and Chong-U Lim (2017). "Reimagining the avatar dream: Modeling social identity in digital media". In: *Communications of the ACM* 60.7, pp. 50–61.

Harrison, Chris, Hrvoje Benko, and Andrew D Wilson (2011). "OmniTouch: wearable multitouch interaction everywhere". In: *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pp. 441–450.

Hawkins, Ian, Muniba Saleem, Bryan Gibson, and Brad J Bushman (2021). "Extensions of the proteus effect on intergroup aggression in the real world." In: *Psychology of Popular Media*.

Heinrich, Chris, Matthew Cook, Tobias Langlotz, and Holger Regenbrecht (2020). "My hands? Importance of personalised virtual hands in a neurorehabilitation scenario". In: *Virtual Reality*, pp. 1–18.

Hildebrandt, Lea K, Cade McCall, and Tania Singer (2017). "Differential effects of attention-, compassion-, and socio-cognitively based mental practices on self-reports of mindfulness and compassion". In: *Mindfulness* 8.6, pp. 1488–1512.

Hilliges, Otmar, David Kim, Shahram Izadi, Malte Weiss, and Andrew Wilson (2012). "HoloDesk: direct 3d interactions with a situated see-through display". In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*. Austin: ACM Press, p. 2421. DOI: 10.1145/2207676.2208405.

Hilsmann, Anna and Peter Eisert (2009). "Tracking and retexturing cloth for real-time virtual clothing applications". In: *International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications*. Springer, pp. 94–105.

Hoang, Thuong N, Hasan S Ferdous, Frank Vetere, and Martin Reinoso (2018). "Body as a Canvas: An Exploration on the Role of the Body as Display of Digital Information". In: *Proceedings of the 2018 Designing Interactive Systems Conference*, pp. 253–263.

Hoermann, Simon, Elizabeth A Franz, and Holger Regenbrecht (2012). "Referred sensations elicited by video-mediated mirroring of hands". In: *PLoS One* 7.12, e50942.

Holle, Henning, Neil McLatchie, Stefanie Maurer, and Jamie Ward (2011). "Proprioceptive drift without illusions of ownership for rotated hands in the "Rubber Hand Illusion" paradigm". In: *Cognitive neuroscience* 2.3-4, pp. 171–178.

Hoyet, Ludovic, Ferran Argelaguet, Corentin Nicole, and Anatole Lécuyer (May 2016). ""Wow! I Have Six Fingers!": Would You Accept Structural Changes of Your Hand in VR?" In: *Frontiers in Robotics and AI* 3. Ed. by Mark Billinghurt. ISSN: 2296-9144. DOI: 10.3389/frobt.2016.00027.

IJsselsteijn, Wijnand A, Yvonne A W de Kort, and Antal Haans (2006). "Is this my hand I see before me? The Rubber Hand Illusion in reality, Virtual Reality, and Mixed Reality". In: *Presence: Teleoperators and Virtual Environments* 15.4, pp. 455–464.

Irigaray, Christian (2017). "Soma Sema: The Body as a Prison for the Soul". In: *Academia.edu* 3, p. 2018.

Ishak, Shaziela, Karen E Adolph, and Grace C Lin (2008). "Perceiving affordances for fitting through apertures." In: *Journal of Experimental Psychology: Human Perception and Performance* 34.6, p. 1501.

Jang, Hyung-Seok, Jae-Yun Jeong, Young-Hyun Kim, Yeo-Jin Yoon, and Sung-Jea Ko (2011). "Augmented Reality with high frame rate for low computational power devices". In: *2011 IEEE International Conference on Consumer Electronics-Berlin (ICCE-Berlin)*. IEEE, pp. 274–275.

Javornik, Ana, Yvonne Rogers, Delia Gander, and Ana Moutinho (2017). "MagicFace: Stepping into character through an Augmented Reality mirror". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 4838–4849.

Jeon, Seokhee and Seungmoon Choi (2009). "Haptic Augmented Reality: Taxonomy and an example of stiffness modulation". In: *Presence: Teleoperators and Virtual Environments* 18.5, pp. 387–408.

Jeunet, Camille, Louis Albert, Ferran Argelaguet, and Anatole Lécuyer (Jan. 2018). ""Do you feel in control?": towards novel approaches to characterise, manipulate and measure the sense of

agency in virtual environments". In: *IEEE transactions on visualization and computer graphics* 24.4, pp. 1486–1495.

Johnson, Adrian S and Yu Sun (2013). "Spatial Augmented Reality on person: Exploring the most personal medium". In: *International Conference on Virtual, Augmented and Mixed Reality*. Springer, pp. 169–174.

Johnson, Kim, Sharron J Lennon, and Nancy Rudd (2014). "Dress, body and self: Research in the social psychology of dress". In: *Fashion and Textiles* 1.1, pp. 1–24.

Kalckert, Andreas and H. Henrik Ehrsson (2012). "Moving A Rubber Hand That Feels Like Your Own: A Dissociation of Ownership and Agency". In: *Frontiers in Human Neuroscience* 6. ISSN: 1662-5161. DOI: 10.3389/fnhum.2012.00040.

Kaneko, Fuminari, Keiichiro Shindo, Masaki Yoneta, Megumi Okawada, Kazuto Akaboshi, and Meigen Liu (2019). "A case series clinical trial of a novel approach using augmented reality that inspires self-body cognition in patients with stroke: effects on motor function and resting-state brain functional connectivity". In: *Frontiers in systems neuroscience* 13, p. 76.

Kato, Hirokazu and Mark Billinghurst (1999). "Marker tracking and hmd calibration for a video-based Augmented Reality conferencing system". In: *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99)*. IEEE, pp. 85–94.

Kato, Hirokazu, Mark Billinghurst, and Ivan Poupyrev (2020). *ARToolKit*. http://github.com/artoolkitx.

Kellems, Ryan O, Cade Charlton, Kjartan Skogly Kversoy, and Miklos Gyori (2020). "Exploring the Use of Virtual Characters (Avatars), Live Animation, and Augmented Reality to Teach Social Skills to Individuals with Autism". In: *Multimodal Technologies and Interaction* 4.3, p. 48.

Kennedy, Robert S, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal (1993). "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness". In: *The international journal of aviation psychology* 3.3, pp. 203–220.

Kilteni, Konstantina, Ilias Bergstrom, and Mel Slater (2013). "Drumming in immersive virtual reality: the body shapes the way we play". In: *IEEE transactions on visualization and computer graphics* 19.4, pp. 597–605.

Kilteni, Konstantina, Raphaela Groten, and Mel Slater (Nov. 2012). "The Sense of Embodiment in Virtual Reality". In: *Presence: Teleoperators and Virtual Environments* 21, pp. 373–387. ISSN: 1054-7460, 1531-3263. DOI: 10.1162/PRES_a_00124.

Kilteni, Konstantina, Antonella Maselli, Konrad P. Kording, and Mel Slater (Mar. 2015). "Over My Fake Body: Body Ownership Illusions For Studying The Multisensory Basis Of Own-body Perception". In: *Frontiers in Human Neuroscience* 9. ISSN: 1662-5161. DOI: 10.3389/fnhum.2015.00141.

Kim, Minseok and Jae Yeol Lee (2016). "Touch and hand gesture-based interactions for directly manipulating 3D virtual objects in

mobile augmented reality". In: *Multimedia Tools and Applications* 75.23, pp. 16529–16550.

Kim, Sanghyun, Junhyung Kim, Ji-Hun Bae, and Jaeheung Park (June 2018). "Real-time Inverse Kinematics Technique for Controlling Redundant Avatar Arm". In.

Kleinberger, Rebecca and Alisha Panjwani (2018). "Digitally Enchanted Wear: a Novel Approach in the Field of Dresses as Dynamic Digital Displays". In: *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 27–34.

Klippel, Alexander, Pejman Sajjadi, Jiayan Zhao, Jan Oliver Wallgrün, Jiawei Huang, and Mahda M Bagher (2021). "Embodied digital twins for environmental applications". In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 4, pp. 193–200.

Kokkinara, Elena and Mel Slater (Jan. 2014). "Measuring the Effects Through Time of the Influence of Visuomotor and Visuotactile Synchronous Stimulation on a Virtual Body Ownership Illusion". In: *Perception* 43, pp. 43–58. ISSN: 0301-0066, 1468-4233. DOI: 10.1068/p7545.

Komiyama, Ryohei, Takashi Miyaki, and Jun Rekimoto (2017). "JackIn Space: Designing a Seamless Transition between First and Third Person View for Effective Telepresence Collaborations". In: *Proceedings of the 8th Augmented Human International Conference*. AH '17. Silicon Valley, California, USA: Association for Computing Machinery. ISBN: 9781450348355. DOI: 10.1145/3041164.3041183. URL: https://doi.org/10.1145/3041164.3041183.

Kraus, Michael W and Wendy Berry Mendes (2014). "Sartorial symbols of social class elicit class-consistent behavioral and physiological responses: A dyadic approach." In: *Journal of Experimental Psychology: General* 143.6, p. 2330.

Kruijff, E., J. E. Swan, and S. Feiner (Oct. 2010). "Perceptual issues in augmented reality revisited". In: *2010 IEEE International Symposium on Mixed and Augmented Reality*, pp. 3–12. DOI: 10.1109/ISMAR.2010.5643530.

Lamounier Jr, EA, K Lopes, A Cardoso, and AB Soares (2012). "Using Augmented Reality techniques to simulate myoelectric upper limb prostheses". In: *J Bioengineer & Biomedical Sci S* 1, p. 2.

Lefcourt, Herbert M (1981). "Differentiating among internality, powerful others, and chance". In: *Research with the locus of control construct* 1. Publisher: Elsevier, pp. 15–59.

Leigh, Sang-won and Pattie Maes (Apr. 2015). "AfterMath: Visualizing Consequences of Actions through Augmented Reality". In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '15. New York, NY, USA: Association for Computing Machinery, pp. 941–946. ISBN: 978-1-4503-3146-3. DOI: 10.1145/2702613.2732695. URL: https://doi.org/10.1145/2702613.2732695.

Lenggenhager, Bigna, Tej Tadi, Thomas Metzinger, and Olaf Blanke (2007). "Video ergo sum: manipulating bodily self-consciousness". In: *Science* 317.5841, pp. 1096–1099.

Likert, Rensis (1932). "A technique for the measurement of attitudes". In: *Archives of psychology* 22. Ed. by R.S. Woodworth.

Lin, Lorraine and Sophie Jörg (July 2016). "Need a hand?: how appearance affects the virtual hand illusion". In: *Proceedings of the ACM Symposium on Applied Perception*. Anaheim California: ACM, pp. 69–76. ISBN: 978-1-4503-4383-1. DOI: 10/gh4tbn. URL: https://dl.acm.org/doi/10.1145/2931002.2931006.

Lincoln, Peter, Alex Blate, Montek Singh, Turner Whitted, Andrei State, Anselmo Lastra, and Henry Fuchs (2016). "From motion to photons in 80 microseconds: Towards minimal latency for virtual and augmented reality". In: *IEEE transactions on visualization and computer graphics* 22.4, pp. 1367–1376.

Lindlbauer, David and Andy D. Wilson (2018). "Remixed Reality: Manipulating Space and Time in Augmented Reality". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. Montreal QC, Canada: ACM Press, pp. 1–13. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173703. URL: http://dl.acm.org/citation.cfm?doid=3173574.3173703.

Lira, Marilia, Julia H. Egito, Patricia A. Dall'Agnol, David M. Amodio, Óscar F. Gonçalves, and Paulo S. Boggio (Nov. 2017). "The influence of skin colour on the experience of ownership in the rubber hand illusion". In: *Scientific Reports* 7.1, p. 15745. ISSN: 2045-2322. DOI: 10/gck739.

Liu, Fei and Stefan Seipel (2017). "On the precision of third person perspective augmented reality for target designation tasks". In: *Multimedia tools and applications* 76.14, pp. 15279–15296.

Lloyd, Donna M. (2007). "Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand". In: *Brain and Cognition* 64.1, pp. 104–109. DOI: https://doi.org/10.1016/j.bandc.2006.09.013.

Longo, Matthew R., Friederike Schüür, Marjolein P.M. Kammers, Manos Tsakiris, and Patrick Haggard (June 2008). "What Is Embodiment? A Psychometric Approach". In: *Cognition* 107, pp. 978–998. ISSN: 00100277. DOI: 10.1016/j.cognition.2007.12.004.

Loper, Matthew, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black (2015). "SMPL: A skinned multi-person linear model". In: *ACM transactions on graphics (TOG)* 34.6, pp. 1–16.

Lueke, Adam and Bryan Gibson (2015). "Mindfulness meditation reduces implicit age and race bias: The role of reduced automaticity of responding". In: *Social Psychological and Personality Science* 6.3, pp. 284–291.

Lugrin, Jean-Luc, Johanna Latt, and Marc Erich Latoschik (Mar. 2015). "Avatar Anthropomorphism and Illusion of Body Ownership in VR". In: *2015 IEEE Conference on Virtual Reality and 3D User*

Interfaces (VR). Arles, France: IEEE, pp. 229–230. ISBN: 978-1-4799-1727-3. DOI: 10.1109/VR.2015.7223379.

Ma, Ke and Bernhard Hommel (2015). "The role of agency for perceived ownership in the virtual hand illusion". In: *Consciousness and cognition* 36, pp. 277–288.

Makhataeva, Zhanat and Huseyin Atakan Varol (2020). "Augmented Reality for Robotics: A Review". In: *Robotics* 9.2, p. 21.

Mark, Leonard S (1987). "Eyeheight-scaled information about affordances: a study of sitting and stair climbing." In: *Journal of experimental psychology: human perception and performance* 13.3, p. 361.

– (2007). "Perceiving the actions of other people". In: *Ecological psychology* 19.2, pp. 107–136.

Mark, Leonard S, James A Balliett, Kent D Craver, Stephen D Douglas, and Teresa Fox (1990). "What an actor must do in order to perceive the affordance for sitting". In: *Ecological Psychology* 2.4, pp. 325–366.

Marotta, Angela, Michele Tinazzi, Clelia Cavedini, Massimiliano Zampini, and Mirta Fiorio (2016). "Individual differences in the rubber hand illusion are related to sensory suggestibility". In: *PloS one* 11.12, e0168489.

Masood, Tariq and Johannes Egger (2019). "Augmented reality in support of Industry 4.0—Implementation challenges and success factors". In: *Robotics and Computer-Integrated Manufacturing* 58, pp. 181–195.

Medeiros, Daniel, Rafael K dos Anjos, Daniel Mendes, João Madeiras Pereira, Alberto Raposo, and Joaquim Jorge (2018). "Keep my head on my shoulders! why third-person is bad for navigation in VR". In: *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10.

Mercier-Ganady, Jonathan, Fabien Lotte, Emilie Loup-Escande, Maud Marchal, and Anatole Lécuyer (2014). "The Mind-Mirror: See your brain in action in your head using EEG and Augmented Reality". In: *2014 IEEE Virtual Reality (VR)*. IEEE, pp. 33–38.

Merleau-Ponty, Maurice (1945). "Phénoménologie de la perception". In: Paris: Gallimard, page 160.

Miyaki, Takashi and Jun Rekimoto (2016). "LiDARMAN: reprogramming reality with egocentric laser depth scanning". In: *ACM SIGGRAPH 2016 Emerging Technologies on - SIGGRAPH '16*. Anaheim, California: ACM Press, pp. 1–2. ISBN: 978-1-4503-4372-5. DOI: 10.1145/2929464.2929481. URL: http://dl.acm.org/citation.cfm?doid=2929464.2929481.

Mori, Masahiro, Karl F. MacDorman, and Norri Kageki (June 2012). "The Uncanny Valley [From the Field]". In: *IEEE Robotics Automation Magazine* 19, pp. 98–100. ISSN: 1558-223X. DOI: 10.1109/MRA.2012.2192811.

Moseley, G Lorimer, Nick Olthof, Annemeike Venema, Sanneke Don, Marijke Wijers, Alberto Gallace, and Charles Spence (2008). "Psychologically induced cooling of a specific body part caused by the

illusory ownership of an artificial counterpart". In: *Proceedings of the National Academy of Sciences* 105.35, pp. 13169–13173.

Muñoz-Salinas, Rafael and Sergio Garrido-Jurado (2014). *ArUco*. https://doi.org/10.1016/j.patcog.2014.01.005.

Nilsen, Trond, Steven Linton, and Julian Looser (2004). "Motivations for augmented reality gaming". In: *Proceedings of FUSE* 4, pp. 86–93.

Nimcharoen, Chontira, Stefanie Zollmann, Jonny Collins, and Holger Regenbrecht (Oct. 2018). ""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)*. Munich, Germany: IEEE, pp. 158–163. ISBN: 978-1-5386-7592-2. DOI: 10.1109/ISMAR-Adjunct.2018.00057.

Nishino, Wataru, Yusuke Yamanoi, Yoshiaki Sakuma, and Ryu Kato (2017). "Development of a myoelectric prosthesis simulator using Augmented Reality". In: *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, pp. 1046–1051.

Noh, Seung-Tak, Hui-Shyong Yeo, and Woontack Woo (2015). "An HMD-based Mixed Reality System for Avatar-Mediated Remote Collaboration with Bare-hand Interaction". In: ed. by P. Figueroa M. Imura and B. Mohler. Kyoto, Japan. DOI: 10.2312/EGVE.20151311.

Numaguchi, Naoki, Atsushi Nakazawa, Takaaki Shiratori, and Jessica K. Hodgins (2011). "A puppet interface for retrieval of motion capture data". en. In: *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation - SCA '11*. Vancouver, British Columbia, Canada: ACM Press, p. 157. ISBN: 978-1-4503-0923-3. DOI: 10.1145/2019406.2019427. URL: http://dl.acm.org/citation.cfm?doid=2019406.2019427.

Oshita, Masaki, Yuta Senju, and Syun Morishige (2013). "Character motion control interface with hand manipulation inspired by puppet mechanism". In: *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, pp. 131–138.

Otono, Riku, Adélaïde Genay, Monica Perusquía-Hernández, Naoya Isoyama, Hideaki Uchiyama, Martin Hachet, Anatole Lécuyer, and Kiyoshi Kiyokawa (Oct. 2022). "Studying "Avatar Transitions" in Augmented Reality: Influence on Sense of Embodiment and Physiological Activity". In: *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. Singapour. DOI: 10.1109/ISMAR-Adjunct57072.2022.00106.

– (Mar. 2023). "Active or Passive Transition? Effects of Avatar Transitions on the Sense of Embodiment and Physiological Activity". In: *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Shanghai, China. URL: https://hal.inria.fr/hal-03968274/.

Parekh, Pranav, Shireen Patel, Nivedita Patel, and Manan Shah (Sept. 2020). "Systematic review and meta-analysis of augmented reality in medicine, retail, and games". In: *Visual Computing for Industry,*

Biomedicine, and Art 3.1, p. 21. ISSN: 2524-4442. DOI: 10/gh635r. URL: https://doi.org/10.1186/s42492-020-00057-7.

Parfit, Derek (1971). "Personal identity". In: *The Philosophical Review* 80.1, pp. 3–27.

Pärtel Lang (2022). *Puppet Master*. Version 1.1. URL: http://root-motion.com/.

Peck, Tabitha C and Mar Gonzalez-Franco (2021). "Avatar embodiment. A standardized questionnaire". In: *Frontiers in Virtual Reality* 1, p. 44.

Peck, Tabitha C, Sofia Seinfeld, Salvatore M Aglioti, and Mel Slater (2013). "Putting yourself in the skin of a black avatar reduces implicit racial bias". In: *Consciousness and cognition* 22.3, pp. 779–787.

Petkova, Valeria I. and H. Henrik Ehrsson (Dec. 2008). "If I Were You: Perceptual Illusion of Body Swapping". In: *PLoS ONE* 3. Ed. by Justin Harris. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0003832.

Petkova, Valeria Ivanova, Mehrnoush Khoshnevis, and H Henrik Ehrsson (2011). "The perspective matters! Multisensory integration in ego-centric reference frames determines full-body ownership". In: *Frontiers in psychology* 2, p. 35.

Piryankova, Ivelina V., Jeanine K. Stefanucci, Javier Romero, Stephan De La Rosa, Michael J. Black, and Betty J. Mohler (Sept. 2014). "Can I Recognize My Body's Weight? The Influence of Shape and Texture on the Perception of Self". In: *ACM Trans. Appl. Percept.* 11.3. ISSN: 1544-3558. DOI: 10.1145/2641568. URL: https://doi.org/10.1145/2641568.

Praetorius, Anna Samira and Daniel Görlich (2020). "How Avatars Influence User Behavior: A Review on the Proteus Effect in Virtual Environments and Video Games". In: *International Conference on the Foundations of Digital Games*. FDG '20. Bugibba, Malta: Association for Computing Machinery. ISBN: 9781450388078. DOI: 10.1145/3402942.3403019. URL: https://doi.org/10.1145/3402942.3403019.

Pujades, Sergi, Betty Mohler, Anne Thaler, Joachim Tesch, Naureen Mahmood, Nikolas Hesse, Heinrich H. Bülthoff, and Michael J. Black (May 2019). "The Virtual Caliper: Rapid Creation of Metrically Accurate Avatars from 3D Measurements". In: *IEEE Transactions on Visualization and Computer Graphics* 25.5. ISSN: 1941-0506. DOI: 10.1109/TVCG.2019.2898748.

Qian, Jing, Jiaju Ma, Xiangyu Li, Benjamin Attal, Haoming Lai, James Tompkin, John F Hughes, and Jeff Huang (2019). "Portal-ble: Intuitive free-hand manipulation in unbounded smartphone-based Augmented Reality". In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 133–145.

Ratan, Rabindra, David Beyea, Benjamin J. Li, and Luis Graciano (2020). "Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: a meta-analysis of the proteus effect". In: *Media Psychology* 23.5, pp. 651–675. DOI: 10.1080/15213269.2019.1623698. eprint: https://doi.org/10.1080/

15213269.2019.1623698. URL: https://doi.org/10.1080/15213269.2019.1623698.

Raz, Liat, Patrice L. Weiss, and Miriam Reiner (2008). "The Virtual Hand Illusion and Body Ownership". In: *Haptics: Perception, Devices and Scenarios*. Ed. by Manuel Ferre. Heidelberg, Germany: Springer, pp. 367–372. ISBN: 978-3-540-69057-3. DOI: 10.1007/978-3-540-69057-3_47.

Reitmayr, Gerhard, Tobias Langlotz, Daniel Wagner, Alessandro Mulloni, Gerhard Schall, Dieter Schmalstieg, and Qi Pan (2010). "Simultaneous localization and mapping for Augmented Reality". In: *2010 International Symposium on Ubiquitous Virtual Reality*. IEEE, pp. 5–8.

Rheu, Minjin, Youjin Jang, and Wei Peng (2020). "Enhancing healthy behaviors through virtual self: A systematic review of health interventions using avatars". In: *Games for health journal* 9.2, pp. 85–94.

Richard, Grégoire, Thomas Pietrzak, Ferran Argelaguet, Anatole Lécuyer, and Géry Casiez (2022). "Within or Between? Comparing Experimental Designs for Virtual Embodiment Studies". In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 186–195. DOI: 10.1109/VR51125.2022.00037.

Rohde, Marieke, Massimiliano Di Luca, and Marc O Ernst (2011). "The Rubber Hand Illusion: feeling of ownership and proprioceptive drift do not go hand in hand". In: *PloS one* 6.6, e21659.

Rolland, Jannick P. and Henry Fuchs (June 2000). "Optical Versus Video See-Through Head-Mounted Displays in Medical Visualization". In: *Presence: Teleoperators and Virtual Environments* 9, pp. 287–309. ISSN: 1054-7460, 1531-3263. DOI: 10.1162/105474600566808.

Roo, Joan Sol and Martin Hachet (Oct. 2017). "One Reality: Augmenting How the Physical World is Experienced by combining Multiple Mixed Reality Modalities". In: *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. Québec City QC Canada: ACM, pp. 787–795. ISBN: 978-1-4503-4981-9. DOI: 10.1145/3126594.3126638. URL: https://dl.acm.org/doi/10.1145/3126594.3126638.

Rosa, Nina (2016). "Player/Avatar Body Relations in Multimodal Augmented Reality Games". In: *Proceedings of the 18th ACM International Conference on Multimodal Interaction - ICMI 2016*. Tokyo, Japan: ACM Press, pp. 550–553. ISBN: 978-1-4503-4556-9. DOI: 10.1145/2993148.2997618.

Rosa, Nina, Jean-Paul van Bommel, Wolfgang Hurst, Tanja Nijboer, Remco C. Veltkamp, and Peter Werkhoven (Mar. 2019). "Embodying an Extra Virtual Body in Augmented Reality". In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Osaka, Japan: IEEE, pp. 1138–1139. ISBN: 978-1-72811-377-7. DOI: 10.1109/VR.2019.8798055.

Rosa, Nina, Remco C. Veltkamp, Wolfgang Hürst, Tanja Nijboer, Carolien Gilbers, and Peter Werkhoven (Aug. 2019). "The Supernumerary Hand Illusion in Augmented Reality". In: *ACM Transac-*

tions on Applied Perception 16, pp. 1–20. ISSN: 15443558. DOI: 10.1145/3341225.

Rossier, J., C. Rigozzi, and S. Berthoud (Mar. 2002). "Validation de la version française de l'échelle de contrôle de Levenson (IPC), influence de variables démographiques et de la personnalité". In: *Annales Médico-psychologiques, revue psychiatrique* 160.2, pp. 138–148. DOI: 10.1016/S0003-4487(01)00111-1.

Roth, Daniel and Marc Erich Latoschik (Dec. 2020). "Construction of the Virtual Embodiment Questionnaire (VEQ)". In: *IEEE Transactions on Visualization and Computer Graphics* 26.12, pp. 3546–3556. ISSN: 1941-0506. DOI: 10/ghfnr2.

Roth, Daniel, Jean-Luc Lugrin, Marc Erich Latoschik, and Stephan Huber (2017). "Alpha IVBO - Construction of a Scale to Measure the Illusion of Virtual Body Ownership". In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. Denver, Colorado, USA: ACM Press, pp. 2875–2883. ISBN: 978-1-4503-4656-6. DOI: 10.1145/3027063.3053272.

Rotter, Julian B. (1966). "Generalized expectancies for internal versus external control of reinforcement." In: *Psychological monographs* 80 1, pp. 1–28.

Saakes, Daniel, Hui-Shyong Yeo, Seung-Tak Noh, Gyeol Han, and Woontack Woo (2016). "Mirror mirror: An on-body t-shirt design system". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 6058–6063.

Sadeghian, Shadan and Marc Hassenzahl (2021). "From Limitations to "Superpowers": A Design Approach to Better Focus on the Possibilities of Virtual Reality to Augment Human Capabilities". In: *Designing Interactive Systems Conference 2021*. DIS '21. Virtual Event, USA: Association for Computing Machinery, pp. 180–189. ISBN: 9781450384766. DOI: 10.1145/3461778.3462111. URL: https://doi.org/10.1145/3461778.3462111.

Salamin, Patrick, Tej Tadi, Olaf Blanke, Frederic Vexo, and Daniel Thalmann (2010). "Quantifying Effects of Exposure to the Third and First-Person Perspectives in Virtual-Reality-Based Training". In: *IEEE Transactions on Learning Technologies* 3.3, pp. 272–276. DOI: 10.1109/TLT.2010.13.

Salm-Hoogstraeten, Sophia von and Jochen Müsseler (May 2021). "Perspective taking while interacting with a self-controlled or independently-acting avatar". In: *Computers in Human Behavior* 118, p. 106698. ISSN: 0747-5632. DOI: 10.1016/j.chb.2021.106698. URL: https://www.sciencedirect.com/science/article/pii/S0747563221000200.

Sanchez, X., Ph. Lambert, G. Jones, and D. J. Llewellyn (2012). "Efficacy of pre-ascent climbing route visual inspection in indoor sport climbing". In: *Scandinavian Journal of Medicine & Science in Sports* 22.1, pp. 67–72. ISSN: 1600-0838. DOI: 10.1111/j.1600-0838.2010.01151.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0838.2010.01151.x.

Sanchez-Vives, Maria V, Bernhard Spanlang, Antonio Frisoli, Massimo Bergamasco, and Mel Slater (2010). "Virtual hand illusion induced by visuomotor correlations". In: *PloS one* 5.4, e10381.

Saxon, Morgan A, Brandon J Thomas, Jeanine K Stefanucci, and Sarah H Creem-Regehr (Sept. 2019). "Does Avatar Presence Facilitate Affordance Judgments from Different Perspectives?" In: *Journal of Vision* 19.10, p. 221c. ISSN: 1534-7362. DOI: 10.1167/19.10.221c. URL: https://doi.org/10.1167/19.10.221c.

Schettler, Aubrie, Vicente Raja, and Michael L. Anderson (Dec. 2019). "The Embodiment of Objects: Review, Analysis, and Future Directions". In: *Frontiers in Neuroscience* 13. Ed. by Sapienza Mariella Pazzaglia, p. 1332. ISSN: 1662-453X. DOI: 10.3389/fnins.2019.01332.

Schoop, Eldon, James Smith, and Bjoern Hartmann (Apr. 2018). "HindSight: Enhancing Spatial Awareness by Sonifying Detected Objects in Real-Time 360-Degree Video". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Montreal QC Canada: ACM, pp. 1–12. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173717. URL: https://dl.acm.org/doi/10.1145/3173574.3173717.

Schultze, Ulrike (2010). "Embodiment and presence in virtual worlds: a review". In: *Journal of Information Technology* 25.4, pp. 434–449.

Schwind, Valentin, Pascal Knierim, Cagri Tasci, Patrick Franczak, Nico Haas, and Niels Henze (2017). ""These are not my hands!": Effect of Gender on the Perception of Avatar Hands in Virtual Reality". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver, Colorado, USA: ACM Press, pp. 1577–1582. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025602.

Seinfeld, Sofia and Jörg Müller (Dec. 2020). "Impact of visuomotor feedback on the embodiment of virtual hands detached from the body". In: *Scientific Reports* 10.1, p. 22427. ISSN: 2045-2322. DOI: 10/gjsnc8. URL: http://www.nature.com/articles/s41598-020-79255-5.

Shibuya, Satoshi, Satoshi Unenaka, and Yukari Ohki (2018). "The relationship between the Virtual Hand Illusion and motor performance". In: *Frontiers in Psychology* 9, p. 2242.

Simons, Daniel J and Christopher F Chabris (1999). "Gorillas in our midst: Sustained inattentional blindness for dynamic events". In: *perception* 28.9, pp. 1059–1074.

Škola, Filip and Fotis Liarokapis (June 2016). "Examining the Effect of Body Ownership in Immersive Virtual and Augmented Reality Environments". In: *The Visual Computer* 32, pp. 761–770. ISSN: 0178-2789, 1432-2315. DOI: 10.1007/s00371-016-1246-8.

Slater, Mel (2004). "How colorful was your day? Why questionnaires cannot assess presence in virtual environments". In: *Presence: Teleoperators & Virtual Environments* 13.4, pp. 484–493.

– (2008). "Towards a Digital Body: The Virtual Arm Illusion". In: *Frontiers in Human Neuroscience* 2. ISSN: 16625161. DOI: 10.3389/neuro.09.006.2008.
– (2018). "Immersion and the illusion of presence in virtual reality". In: *British Journal of Psychology* 109.3, pp. 431–433. DOI: https://doi.org/10.1111/bjop.12305.

Slater, Mel, Daniel Pérez Marcos, Henrik Ehrsson, and Maria V Sanchez-Vives (2008). "Towards a digital body: the virtual arm illusion". In: *Frontiers in human neuroscience* 2, p. 6.

Slater, Mel, Bernhard Spanlang, Maria V. Sanchez-Vives, and Olaf Blanke (May 2010). "First Person Experience of Body Transfer in Virtual Reality". In: *PLoS ONE* 5. Ed. by Mark A. Williams. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0010564.

Slepian, Michael L, Simon N Ferber, Joshua M Gold, and Abraham M Rutchick (2015). "The cognitive consequences of formal clothing". In: *Social Psychological and Personality Science* 6.6, pp. 661–668.

Spanlang, Bernhard, Jean-Marie Normand, David Borland, Konstantina Kilteni, Elias Giannopoulos, Ausiàs Pomés, Mar Gonzalez-Franco, Daniel Perez-Marcos, Jorge Arroyo-Palacios, Xavi Navarro Muncunill, et al. (2014). "How to build an embodiment lab: achieving body representation illusions in Virtual Reality". In: *Frontiers in Robotics and AI* 1, p. 9.

Stone, Gregory P (1962). "Appearance and the self". In: *Human behavior and social processes*, pp. 86–118.

Sutherland, Ivan (1965). "The ultimate display". In.

Sutherland, Ivan E (1968). "A head-mounted three dimensional display". In: *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, pp. 757–764.

Syiem, Brandon Victor, Ryan M. Kelly, Jorge Goncalves, Eduardo Velloso, and Tilman Dingler (2021). "Impact of Task on Attentional Tunneling in Handheld Augmented Reality". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan: Association for Computing Machinery. ISBN: 9781450380966. DOI: 10.1145/3411764.3445580. URL: https://doi.org/10.1145/3411764.3445580.

Teräs, Marko (July 2015). *You Are Your Avatar Is You: Phenomenological Literature Review of Virtual Embodiment in Virtual Environments*. Tech. rep.

Thomas, Bruce H (2012). "A survey of visual, mixed, and augmented reality gaming". In: *Computers in Entertainment (CIE)* 10.1, pp. 1–33.

Toothman, Nicholas and Michael Neff (Mar. 2019). "The Impact of Avatar Tracking Errors on User Experience in VR". In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Osaka, Japan: IEEE, pp. 756–766. ISBN: 978-1-72811-377-7. DOI: 10.1109/VR.2019.8798108.

Treepong, Bantita, Hironori Mitake, and Shoichi Hasegawa (Nov. 2018). "Makeup Creativity Enhancement with an Augmented Reality Face Makeup System". In: *Computers in Entertainment* 16.4,

6:1–6:17. DOI: 10.1145/3277452. URL: https://doi.org/10.1145/3277452.

Tsakiris, Manos and Patrick Haggard (2005). "The rubber hand illusion revisited: visuotactile integration and self-attribution". In: *Science*.

Tsakiris, Manos, Gita Prabhu, and Patrick Haggard (June 2006). "Having A Body Versus Moving Your Body: How Agency Structures Body-Ownership". In: *Consciousness and Cognition* 15, pp. 423–432. ISSN: 10538100. DOI: 10.1016/j.concog.2005.09.004.

Walsh, Eamonn, DN. Guilmette, Matthew R. Longo, James W. Moore, DA. Oakley, PW. Halligan, MA. Mehta, and Q. Deeley (June 2015). "Are You Suggesting That's My Hand? The Relation Between Hypnotic Suggestibility and the Rubber Hand Illusion". In: *Perception* 44, pp. 709–723. ISSN: 0301-0066, 1468-4233. DOI: 10.1177/0301006615594266.

Waltemate, Thomas, Dominik Gall, Daniel Roth, Mario Botsch, and Marc Erich Latoschik (2018). "The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response". In: *IEEE transactions on visualization and computer graphics* 24.4, pp. 1643–1652.

Waltemate, Thomas, Irene Senna, Felix Hülsmann, Marieke Rohde, Stefan Kopp, Marc Ernst, and Mario Botsch (2016). "The Impact of Latency on Perceptual Judgments and Motor Performance in Closed-loop Interaction in Virtual Reality". In: *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology - VRST '16*. Munich, Germany: ACM Press, pp. 27–35. ISBN: 978-1-4503-4491-3. DOI: 10.1145/2993369.2993381.

Wang, Isaac, Jesse Smith, and Jaime Ruiz (May 2019). "Exploring Virtual Agents for Augmented Reality". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow Scotland Uk: ACM, pp. 1–12. ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300511.

Wang, Kai, Daisuke Iwai, and Kosuke Sato (2017). "Supporting Trembling Hand Typing Using Optical See-Through Mixed Reality". In: *IEEE Access* 5, pp. 10700–10708. ISSN: 2169-3536. DOI: 10/ghmtd3.

Wang, Tzu-Yang, Yuji Sato, Mai Otsuki, Hideaki Kuzuoka, and Yusuke Suzuki (Mar. 2019). "Effect of Full Body Avatar in Augmented Reality Remote Collaboration". In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Osaka, Japan: IEEE, pp. 1221–1222. ISBN: 978-1-72811-377-7. DOI: 10.1109/VR.2019.8798044.

Wang, Yuntao, Zhou Su, Ning Zhang, Dongxiao Liu, Rui Xing, Tom H Luan, and Xuemin Shen (2022). "A survey on metaverse: Fundamentals, security, and privacy". In: *arXiv preprint arXiv:2203.02662*.

Warren Jr, William H and Suzanne Whang (1987). "Visual guidance of walking through apertures: body-scaled information for affordances." In: *Journal of experimental psychology: human perception and performance* 13.3, p. 371.

Whitebread, David, Marisol Basilio, Martina Kuvalja, and Mohini Verma (2012). "The importance of play". In: *University of Cambridge, Toy Industries of Europe. The importance of play*.

Wolf, Erik, Nina Döllinger, David Mal, Carolin Wienrich, Mario Botsch, and Marc Erich Latoschik (2020). "Body Weight Perception of Females using Photorealistic Avatars in Virtual and Augmented Reality". In: *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, pp. 462–473.

Wolf, Erik, Marie Luisa Fiedler, Nina Döllinger, Carolin Wienrich, and Marc Erich Latoschik (2022). "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. DOI: 10.1109/VR51125.2022.00054.

Yee, Nick and Jeremy Bailenson (July 2007). "The Proteus Effect: The Effect of Transformed Self-Representation on Behavior". In: *Human Communication Research* 33, pp. 271–290. ISSN: 0360-3989, 1468-2958. DOI: 10.1111/j.1468-2958.2007.00299.x.

Yim, Mark Yi-Cheon and Sun-Young Park (July 2019). ""I am not satisfied with my body, so I like augmented reality (AR)": Consumer responses to AR-based product presentations". In: *Journal of Business Research* 100, pp. 581–589. ISSN: 0148-2963. DOI: 10/gh7x7p. URL: https://www.sciencedirect.com/science/article/pii/S0148296318305228.

Yoon, Gunwoo and Patrick T Vargas (2014). "Know thy avatar: The unintended effect of virtual-self representation on behavior". In: *Psychological science* 25.4, pp. 1043–1045.

Zabels, Roberts, Rendijs Smukulis, Ralfs Fenuks, Andris Kučiks, Elza Linina, Kriss Osmanis, and Ilmars Osmanis (2021). "Reducing motion to photon latency in multi-focal augmented reality near-eye display". In: *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) II*. Vol. 11765. International Society for Optics and Photonics, 117650W.

Zheng, Feng, Turner Whitted, Anselmo Lastra, Peter Lincoln, Andrei State, Andrew Maimone, and Henry Fuchs (2014). "Minimizing latency for Augmented Reality displays: Frames considered harmful". In: *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, pp. 195–200.

Zollmann, Stefanie, Raphael Grasset, Tobias Langlotz, Wei Hong Lo, Shohei Mori, and Holger Regenbrecht (2020). "Visualization Techniques in Augmented Reality: A Taxonomy, Methods and Patterns". In: *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1. ISSN: 1941-0506. DOI: 10/gh52t6.

Zöllner, Michael, Stephan Huber, Hans-Christian Jetter, and Harald Reiterer (2011). "NAVI – A Proof-of-Concept of a Mobile Navigational Aid for Visually Impaired Based on the Microsoft Kinect". In: *Human-Computer Interaction – INTERACT 2011*. Ed. by David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz,

C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi, Gerhard Weikum, Pedro Campos, Nicholas Graham, Joaquim Jorge, Nuno Nunes, Philippe Palanque, and Marco Winckler. Vol. 6949. Series Title: Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 584–587. ISBN: 978-3-642-23767-6 978-3-642-23768-3. DOI: 10.1007/978-3-642-23768-3_88. URL: http://link.springer.com/10.1007/978-3-642-23768-3_88.

Zoran, Amit and Joseph A. Paradiso (Apr. 2013). "FreeD: a freehand digital sculpting tool". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Paris France: ACM, pp. 2613–2616. ISBN: 978-1-4503-1899-0. DOI: 10.1145/2470654.2481361. URL: https://dl.acm.org/doi/10.1145/2470654.2481361.

Appendix A
Experimental material of Chapter 5

A.1 Yes-or-no and estimation questions

Table A.1: The evaluation questions we asked orally after each experimental condition (translated from French). For yes-or-no questions (YG1-7 and YB1-4), participants also provided a level of confidence on a scale from 1 to 10. For estimation questions, (EG1-5 and EB1-4), they answered with a unit on a scale projected in the environment and with a margin of error (e.g. "21 plus or minus 2").

Environment	ID	Statement
Ground		"*On the sheet-covered platform, do you think...*
	YG1	"You could walk under the chain suspended on the left without touching it, without bending your head?"
	YG2	"Your heels could reach the block on the left if you lied down with your head against the block on the right?"
	YG3	"You could touch the chain suspended on the left if you were standing in the middle of the left block? (facing us, without bending over)"
	YG4	"You could touch with your hands the [*color A*] grip on the left and the [*color B*] grip on the right at the same time if you stood in the middle?"
	YG5	"You could touch with your hands the [*color A'*] grip on the left and the [*color B'*] grip on the right at the same time if you stood in the middle?"
	YN6	"You could touch the chain suspended on the left if you were standing underneath it and stretched your arms above your head, without standing on your tiptoes?"
	YG7	"You could touch the black beam above if you were standing on the left block and stretched your arms above your head? (without standing on your tiptoes)"
		... With what level of confidence from 1 to 10?"
		"*On the arbitrary scale projected on the wall, which unit do you think...*
	EG1	"The top of your head would reach if you were sitting on the block on the left?"
	EG2	"The top of your head would reach if you were standing on the block on the right?"
	EG3	"Your heels would reach if you lied down with your head against the block on the left?"
	EG4	"Your hand would reach if you extended your arms above your head while standing in the middle?"
	EG5	"Your heels would reach if you stretched your legs toward the right while sitting against the block on the left?"
		... With what margin of error?"
Beam		"*On the black beam up there, do you think...*
	YB1	"Your hands could reach the groove on the wall behind if you stretched your arms above your head? (without standing on your tiptoes)"
	YB2	"Your heels could reach the tape on the left if you lie down with your head against the suspension cable on the left?"
	YB3	"Your hands could reach the suspension cable on the far right if you stretched your arms towards it while lying down, with your heels located against the tape on the left?"
	YB4	"Your hand could touch the suspension cable on the left if you were standing on the tape on the left? (facing us, without bending over)"
		... With what level of confidence from 1 to 10?"
		"*On the arbitrary scale projected on the wall, which unit do you think...*
	EB1	"The top of your head would reach if you were standing on the beam?"
	EB2	"The top of your head would reach if you were sitting on the beam?"
	EB3	"Your hand would reach if you extended it horizontally towards the right while facing us, with the middle of your body aligned on the red line?"
	EB4	"Your heels would reach (on the left) if you lied down on the beam with your head against the suspension cable on the far right?"
		... With what margin of error?"

A.2 Post-experiment questionnaire of Exp. 1

Table A.2: Subjective questionnaire results of the Experiment 1 presented in Chapter 4 (items translated from French. Participants answered the statements Q1-11 on a 7-point Likert scale. Items C1-C5 were optional open-ended text boxes allowing participants to write their points of view. All items were in common with the questionnaire of Experiment 2, except for items Q7, Q8, and Q9 (marked with a star).

Subject	ID	Statement	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree
Appearance	Q1	"My avatar's morphology was similar to mine."	0%	0%	12%	0%	31%	50%	6%
	Q2	"I found the appearance of my avatar disturbing or distracting."	50%	38%	0%	6%	6%	0%	0%
	C1	*Comments on the appearance of the avatar (optional)*							
Presence	Q3	"My avatar seemed to be present with me in the real world."	0%	19%	6%	6%	25%	31%	12%
	Q4	"I had the feeling that my avatar was not well integrated into the environment."	12%	12%	31%	6%	19%	12%	6%
	C2	*Comments on the presence/integration of the avatar.*							
Control	Q5	"I managed to put my avatar in the positions I wanted."	0%	0%	6%	0%	6%	50%	38%
	Q6	"In general, I found it difficult to control my avatar."	19%	50%	0%	6%	19%	6%	0%
	C3	*Comments on the control of the avatar.*							
Help	Q7*	"Seeing my avatar helped me respond with more confidence."	0%	0%	12%	0%	0%	12%	75%
	Q8*	"I think my avatar correctly represents the space my body would occupy where I placed it."	0%	0%	25%	12%	0%	38%	25%
	Q9*	"I think my avatar may have misled me when I answered the evaluation questions."	12%	50%	0%	12%	12%	12%	0%
	Q10	"Seeing my avatar made it difficult for me to answer the questions."	75%	12%	12%	0%	0%	0%	0%
	C4	*Comments on the help provided by the avatar.*							
Enjoyment	Q11	"I enjoyed using my avatar."	0%	6%	0%	6%	0%	44%	44%
	C5	*Comments on the appreciation of the avatar.*							

A.3 Post-experiment questionnaire of Exp. 2

Table A.3: Subjective questionnaire results of the Experiment 2 presented in Chapter 4 (items translated from French). Participants answered the statements Q1-11 on a 7-point Likert scale. Items C1-C5 were optional open-ended text boxes allowing subjects to write their points of view. All items were in common with the questionnaire of Experiment 1, except for item Q7 (marked with a star).

Subject	ID	Statement	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree
Appearance	Q1	"My avatar's morphology was similar to mine."	0%	0%	13%	0%	20%	47%	20%
	Q2	"I found the appearance of my avatar disturbing or distracting."	47%	27%	0%	13%	13%	0%	0%
	C1	*Comments on the appearance of the avatar (optional)*							
Presence	Q3	"My avatar seemed to be present with me in the real world."	0%	13%	7%	7%	47%	20%	7%
	Q4	"I had the feeling that my avatar was not well integrated in the environment."	0%	40%	20%	0%	20%	13%	7%
	C2	*Comments on the presence/integration of the avatar (optional)*							
Control	Q5	"I managed to put my avatar in the positions I wanted."	0%	0%	13%	7%	0%	53%	27%
	Q6	"In general, I found it difficult to control my avatar."	13%	47%	13%	0%	20%	7%	0%
	C3	*Comments on the control of the avatar (optional)*							
Help	Q7*	"Seeing my avatar helped me respond with more confidence than in the condition where I did not use my avatar."	0%	7%	13%	0%	0%	53%	0%
	Q10	"Seeing my avatar made it difficult for me to answer the questions."	53%	33%	13%	0%	0%	0%	0%
	C4	*Comments on the help provided by the avatar (optional)*							
Enjoyment	Q11	"I enjoyed using my avatar."	0%	0%	7%	0%	13%	60%	20%
	C5	*Comments on the appreciation of the avatar (optional)*							

A.4 Post-experiment questionnaire of Exp.3

Table A.4: Subjective questionnaire results of the Experiment 3 presented in Chapter 4 (items translated from French). Participants answered the statements Q1-14 on a 5-point Likert scale. Items C1-C5 were optional open-ended text boxes allowing subjects to write their points of view.

Subject	ID	Statement	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Attention	Q1	"I was able to focus on the task rather than the headset and virtual display."	0%	0%	0%	28%	72%
	Q2	"I got used to the display of the virtual content."	0%	0%	0%	48%	52%
	Q3	"The quality of the visual display made it difficult or distracting for me to answer the questions."	57%	29%	14%	0%	0%
	C1	*Comments on the display (optional)*					
Door	Q4	"The virtual door seemed to be present with me in the real world."	10%	24%	33%	19%	14%
	Q5	"I felt that the virtual door was not well integrated into the environment."	14%	38%	14%	33%	0%
	C2	*Comments on the door (optional)*					
Avatar	Q6	"In general, I found it difficult to control my avatar."	32%	27%	27%	14%	0%
	Q7	"My avatar's morphology was similar to mine."	5%	0%	18%	55%	23%
	Q8	"I found the appearance of my avatar disturbing or distracting."	64%	27%	9%	0%	0%
	C3	*Comments on the control and appearance of the avatar (optional)*					
Help	Q9	"I could easily estimate my ability to get through the door during the experiment.	0%	2%	40%	57%	0%
	Q10	"Using my avatar helped me estimate my ability to get through the door."	0%	9%	18%	36%	36%
	Q11	"I used my training with the avatar to answer the questions."	0%	9%	5%	59%	27%
	Q12	"Using my avatar allowed me to answer questions with more confidence."	0%	5%	18%	45%	32%
	Q13	"I think the avatar may have misled me when I answered the questions."	50%	32%	9%	9%	0%
	C4	*Comments on the help provided by the avatar (optional)*					
Enjoyment	Q14	"I enjoyed using my avatar."	0%	0%	5%	45%	50%
	C5	*Comments on the appreciation of the avatar (optional)*					

A.5 Wariness of new technologies scale

Table A.5: Questionnaire applied before Experiment 3 to collect demographic data (D1-5) and evaluate the user's opinion on new technologies (W1-5). We used the "Wariness of New Technologies Scale" of Anania (2019), translated to French. Participants answered W1-5 on a 5-point Likert scale.

ID	Statement
D1	"Age"
D2	"Gender"
D3	"Body height"
D4	"How many hours do you spend playing video games per week?"
D5	"Have you used a virtual or augmented reality headset before?"

		Strongly disagree	Disagree	Neutral	Agree	Strongly agree
W1	"In general, I am wary of new technology."	33%	50%	14%	2%	0%
W2	"New technology scares me."	64%	31%	2%	2%	0%
W3	"New technology is not as safe as it should be."	24%	31%	31%	12%	2%
W4	"I tend to fear new technology until it is proven to be safe."	21%	48%	14%	14%	2%
W5	"New technology is likely to be dangerous."	7%	17%	31%	33%	12%

Appendix B
Experimental material of Chapter 6

B.1 Mindfulness instructions

The following instructions were given by the experimenter:

> Please close your eyes. I will ask you to imagine yourself transforming into the Hulk. First, focus on your breath. You are becoming conscious of your body. You can feel your heartbeat. You can feel the sensations on your skin, of your muscles. You feel very sensitive to everything. Now, try to think of something that made you stressed or angry recently. If nothing comes up, try to imagine someone attempting to harm a person you love, in front of you. You have started to transform into Hulk. You can feel your body change: your skin is progressively turning green. You are becoming bigger, larger. You feel your muscles becoming very tense, rigid. You start feeling powerful, like you could break anything. Keep concentrating on your anger. The more you think of it, the stronger you feel, and the more your body changes. *Continues like this repeating variations until 5 minutes are reached.*

B.2 Acting instructions

The following instructions were given by the experimenter:

> Please open your eyes, stand up, and walk to the middle of the room. You are now Hulk. Try to act like him by following my instructions. You may also improvise freely if you want. First, you can look at your body and try to picture your skin being green, your limbs being much larger. Your muscles seem to be pulsing. Try to contract them. Imagine your heartbeat accelerating. Breathe deeply, slowly. Continue to recall the moment that made you angry. Clench your fingers into a tight fist. Now, arch your body. Take an aggressive posture, as if preparing to fight. You are extremely angry now, and it shows on your face. You feel like you want to scream. You feel like you want to punch the walls and kick the things around you. But instead, you just walk around with heavy steps, to keep control of yourself. Start walking like your body is very large. *Continues like this repeating variations until 5 minutes are reached.*

B.3 Virtual Embodiment Questionnaire (VEQ)

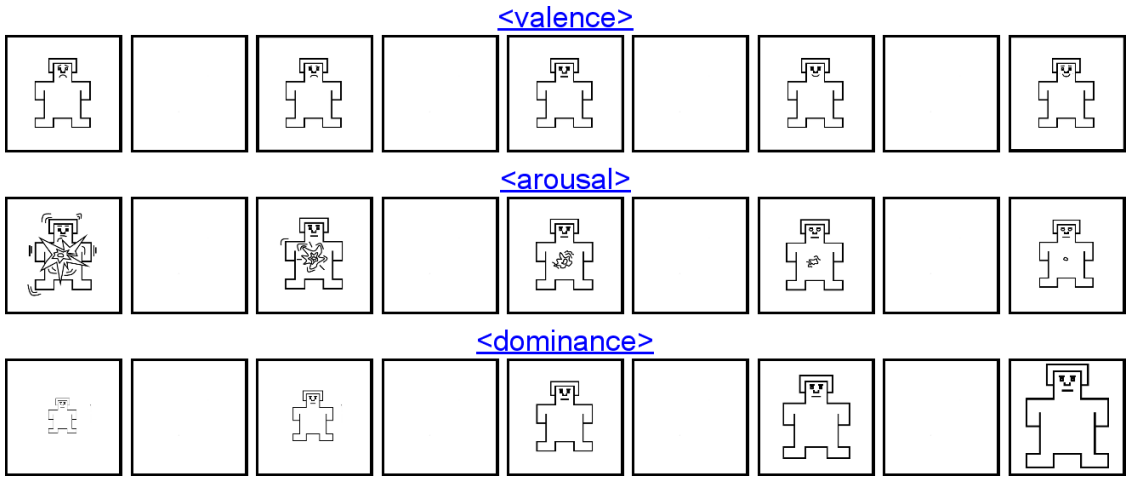
Table B.3: The embodiment questionnaire, adapted from Peck and Gonzalez-Franco (2021).

We report the median and the first and third quartiles for each item. The "Score" column gives the contribution of each item to the following scores: Appearance (A), Response (R), Ownership (O), and Multisensory (M).

ID	Score	Questions	Condition A	Condition B
Q1	A	"I felt out of my body."	3[2, 4.5]	3.5[2, 6]
Q2	A	"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."	4[3, 5]	5[3.5, 6]
Q3	A,M	"I felt as if the movements of the virtual body were influencing my own movements."	4.5[3.75, 6]	4.5[3, 6]
Q4	A,R	"It felt as if my (real) body were turning into an "avatar" body."	5.5[3.75, 6]	5.5[3.75, 6.5]
Q5	A,O	"It felt as if my real body was starting to take on the posture or shape of the virtual body that I saw."	5.5[3.75, 6]	6[4, 6.25]
Q6	A,R	"I felt like I was wearing different clothes from when I came to the laboratory."	3.5[2, 2.5]	4[2.75, 5]
Q7	R	"I felt as if my body was stronger."	5[3, 6]	6[5, 6.25]
Q8	R	"I felt a realistic sensation in my body when I broke the stones."	5[4, 6]	6[5, 6]
Q9	A,R	"I felt that my own body could have been affected by the virtual world."	5[2.75, 6]	5[3.75, 5]
Q10	O	"I felt as if the virtual body was my body. "	5[3, 6]	5.5[3.75, 6]
Q11	O	"It felt that the virtual body resembled my own (real) body in terms of shape/skin tone/other visual features. "	4[2, 5]	4.5[3.75, 6]
Q12	O,M	"I felt as if my body was located where I saw the virtual body"	5.5[4, 6]	6[5.75, 6]
Q13	O,M	"I felt like I could control the virtual body as if it was my own body."	5[5, 6]	6[6, 7]
Q14	O,M	"It seemed as if I felt the touch of the stones in the location where I saw the virtual hands touched."	5[4, 5]	6[5, 6]
Q15	R,M	"It seemed as if the touch I felt was caused by the stones touching the virtual hands. "	5[4, 5.25]	5[4.75, 5.25]
Q16	A,M	"It seemed as if my feet were touching the virtual floor."	5[4, 5.25]	6[5.75, 7]

B.4 Self-Assessment Manikin (SAM)

Figure B.4: Self-Assessment Manikin scale (Bradley et al., 1994) applied before, during, and after the embodiment of the Hulk avatar. Users had to rate the valence, arousal, and dominance of their emotional state by selecting the corresponding manikin or box. The answers were on a 9-point Likert scale.



Acronyms

1PP First-person Perspective. 24–26, 30, 34, 39, 91, 106, 109, 110, 120

3PP Third-person Perspective. 24–26, 34, 35, 39, 61, 65, 66, 69, 71, 73, 75, 76, 83, 84, 91, 99, 114, 120

AR Augmented Reality. 9

AV Augmented Virtuality. 10

BCI Brain-computer Interface. 123

BOI Body Ownership Illusion. 11, 17, 18, 23, 25, 27, 31–34, 39, 40, 43, 51, 52, 54, 57, 58, 104–107, 118–120

CAVE Cave Automatic Virtual Environment. 9, 26

EEG Electroencephalography. 32, 56

HDD Head-down Display. 89

HMD Head-mounted Display. 26, 27, 67, 68, 78, 83, 85, 120, 123, 124

HUD Head-up Display. 89

IK Inverse Kinematics. 27, 30, 70, 108

IoT Internet of things. 23

LoC Locus of Control. 42, 44, 46, 47, 54

MR Mixed Reality. 10, 11, 63, 65–67, 71, 99

MRTK Mixed Reality Tool Kit. 24, 69, 70

OST Optical See-Through. 26, 28, 30, 37–42, 44, 51, 52, 57, 58, 66–68, 118, 120

RHI Rubber Hand Illusion. 11, 56

SAM Self-Assessment Manikin. 110, 111

SCR Skin Conductance Response. 32

sEMG Surface Electromyography. 114

SLAM Simultaneous Localization and Mapping. 29

SoE Sense of Embodiment. 11, 12, 14, 17, 24, 25, 27, 28, 31–35, 38–41, 43–45, 47, 49–55, 57, 58, 99, 103–105, 107, 109, 111–114, 117–121

VHI Virtual Hand Illusion. 32, 33, 39, 43

VM Visuomotor. 33, 40, 41, 43, 109, 110

VR Virtual Reality. 8, 41, 66

VST Video See-Through. 25, 26, 28, 39, 40, 120

VT Visuotactile. 33, 34, 40, 105

WoW Windows-on-the-world. 25

XR Extended Reality. 9, 100, 122