Supernumerary Embodiment in Multimodal Augmented Reality Games

The author of this dissertation received financial support from the Netherlands Organisation for Scientific Research (NWO) for the research conducted within the Game Research Graduate Program (project number 022.005.017).

Cover art: Maaike Spelt
Cover design: Maaike Spelt
Lay-out: Nina Rosa
Printed by: Duurzame Dissertatie

The pages of this dissertation were printed on 100% post-consumer recycled paper which carries the Blaue Engel certification label. The cover was printed on Paperwise® paper, which is made from agricultural waste. The printing office is ISO 14001 certified.

A tree has been planted for every copy of this thesis.

ISBN: 978-90-393-7420-7
URL: `http://www.nerosa.nl/dissertation/`

Supernumerary Embodiment in Multimodal Augmented Reality Games

Boventallige Belichaming in Multimodale Augmented Reality Games

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 22 november 2021 des middags te 2.15 uur

door

Nina Emmanuelle Rosa

geboren op 9 oktober 1992
te Elizabeth, New Jersey, Verenigde Staten van Amerika

Promotoren:

Prof. dr. P.J. Werkhoven
Prof. dr. R.C. Veltkamp

Copromotor:

Dr. W.O. Hürst

Contents

Chapter 1

Introduction

Gaming has changed rapidly over the last few decades. The earliest video games were very simple, and branched off of simulations. Besides the technological advances which moved playing to screens and handheld controllers, the scope of the games changed as well. The games were given extensive storylines and allowed players to explore new virtual worlds. The advances in virtual reality (VR) technology were also visible in the emergence of 3D games. Now both VR and augmented reality (AR) gaming have become a reality. Using handheld devices and/or wearables, players can experience new worlds with their whole body: they can see the virtual or merged real and virtual world from a first person perspective (1PP), touch virtual objects with touch-feedback, and interact with objects using their own bodies through motion tracking.

VR games can already be found in many households due to the technology becoming smaller and more affordable. AR gaming on the other hand is still in its infancy. AR gaming can generally refer to many types of gaming methods, such as location-based AR where the real locations are augmented with virtual information, or perceptual AR where the real senses are augmented with virtual information. In this dissertation I shall only focus on the latter: specifically, immersive AR where the user is placed in the merged world from a correct perspective using a head-mounted display (HMD). The required technology for merging real and virtual sensory information is still too complex and expensive for arcade or home gaming, and there are only a handful of commercial titles. Nonetheless, the AR gaming market is expected to grow steadily in the next few years.

Gaming technologies have changed not only the way we see and control games, but also how we experience them on a higher level. We feel connected to the game. We may be represented by a virtual avatar, and by using VR and AR HMDs we can also now place ourselves in the game perceptually, to a greater extent than 1PP and third-person perspective (3PP) games on traditional screens allowed, and experience the game world in a multisensory way. This creates a completely different gaming experience than games from decades ago. We experience storylines and explore worlds first hand, and act in these worlds as if they were real. In the context of entertainment games, this may offer a player the opportunity to briefly let go of stressful daily life and experience a second life. On the other hand, such almost real experiences offer the opportunity to simulate extreme real world scenarios for, for example, military training purposes, for which the technology was originally developed, but also in the areas of medicine, maintenance, and telerobotics.

A reason AR is not yet as prominent in these domains as VR is partially due to its technological complexity in terms of correctly mixing real and virtual and correctly presenting consistent multisensory feedback. If we compare purely visual VR with purely visual AR, we see that in both cases the virtual elements must be presented in a logical manner to the user, but in the case of AR the virtual elements must also cohere with the current real surroundings. This requires a multitude of cameras, some integrated into the HMD, together with heavy behind the scenes calculations. Adding another sense such as audio or touch increases this complexity drastically. Despite these obstacles, more and more studies are proving these multimodal interactions in AR will be possible in the near future. One could say that a goal of these technological developments is to create such a well-merged scene, that the user can no longer differentiate between what is real and what is virtual. In VR we have prior knowledge that the entire scene is fake, and in AR we have prior knowledge about which components are real, meaning everything else must be fake. Interestingly, there are studies outside of AR that would suggest we can accept acknowledged fake objects as real, namely with respect to one's body.

Both entertainment and serious applications of virtual technology rely on something games can do very well: they connect the player to their in-game avatar. Players tend to feel as if they are 'in' the game, even in traditional screen-controller setups. The views on what exactly this player-avatar link is vary, some stating that it can change at any moment depending on the type of game play, other stating that it in many ways resembles how we feel about our real bodies in the real world, that is, embodiment.

In the real world, although we are not always aware of it, our body is constantly processing external sensory signals that together tell us something about what is our body, what it is doing, and where it is in space. These three components, body ownership, agency, and self-location, respectively, together form the sense of embodiment (Kilteni et al., 2012a). Studies on real world embodiment are numerous, and typically include a variant of the Rubber Hand Illusion (RHI) (Botvinick and Cohen, 1998). In such an illusion, participants experience a fake hand as their own after receiving synchronous tactile stimulation of the real and rubber arm. Variations of this illusion have also used other body parts or entire bodies, and have been performed numerously in reality and VR, but infrequently in AR.

Studies on embodiment as a player-avatar link tend to focus on aspects of the gaming setup such as viewing perspective (1PP versus 3PP), multimodal feedback (e.g. with versus without controller vibrations), and form of interaction (e.g. controller versus motion capture), and how these may affect embodiment and related concepts such as presence. Results typically indicate that these concepts are more strongly experienced when the game allows for a more natural setting; that is, when the game interaction resembles real world interaction, players tend to experience bodily concepts known from the real world. For VR embodiment in particular, studies repeatedly find the same necessary factors.

For AR embodiment, determining the necessary factors becomes complicated. Certainly a realistic scene seems favorable, but what does realistic even mean in an AR sense? One can always try to increase the quality of various sensory signals, but as illustrated earlier, technology is not yet at a stage where multiple modalities are easily displayed in AR. In the early stages of VR there was a prominent philosophical discussion of whether the user may ever forget about reality and take on the virtual environment as the real environment, which explained why users would react in certain ways: screaming when a monster approaches, being afraid of falling off a virtual cliff, trying to lean on virtual objects that are not there. But in AR users are constantly reminded of the real world, because they continue to experience it together with the virtual components. It would then seem that it becomes more difficult to experience a virtual object as real, since users are constantly reminded that it is fake. Taking this back to the context of an RHI, in the experiments that take place in reality, subjects are wondering whether a fake hand could be their real hand, but there is no question of whether the fake hand factually exists; a virtual hand illusion in AR on the other hand becomes much more difficult to achieve, because the virtual hand is not only fake, but also does not even really exist. The few studies on body ownership in AR have shown that the medium still allows the experience of ownership of a virtual arm (Suzuki et al., 2013), although it may weaken the experience compared to reality (Škola and Liarokapis, 2016) and VR (IJsselsteijn et al., 2006).

This would suggest that any embodied experience in AR may require a greater degree of willingness by the user to accept the scene as realistic than in other media.

Although there has been a surge in literature on AR technological developments, an understanding of experiences in AR is currently lacking. This is despite a great relevance of, for example, achieving a sense of embodiment in AR, as it is believed to be related to other application goals and could therefore provide not only experiential benefits but also practical benefits. For example, prior research has shown that factors such as embodiment and task effectiveness in teleoperation tasks are somehow positively correlated, but have been rarely systematically modelled in reality and VR, let alone in AR. With the recent advances in HMDs for AR, there is also an inclination towards a 1PP setups where the user interacts with virtual objects using the real hands. However, once the user wants to interact with far away objects, a form of gestural interaction is required, resulting in a divided interaction experience. Adding a third virtual hand that interacts in a similar fashion as the real hand and operates at a larger distance than the boundary of reaching space may amend this. Then, the question remains whether the previously found positive correlation between embodiment and performance in other media also exists in AR.

1.1 Dissertation Overview

The focus of this dissertation is the existence and function of embodiment of supernumerary virtual body parts alongside the real body in multimodal augmented reality games. This is investigated in four steps, each corresponding to a chapter, where the first two chapters create a foundation for the last two chapters. In Chapter 2, I first take a step back and discuss what AR is, and how multimodality expands its scope, that is, how certain seemingly not-AR multimodal applications should in fact be categorized as AR. I redefine AR in terms of different types of stimuli, and introduce a new classification system for multimodal AR, as existing AR classification systems are not well defined for multimodality.

Chapter 3 introduces the avatar as a body representation in games and discusses player-avatar links from an interdisciplinary perspective. Specifically, I outline phenomenological accounts of player embodiment in games, player-avatar links beyond embodiment, the state of empirical embodiment research in reality and VR, and both objective and subjective qualities that may influence concepts similar to embodiment in games. I highlight issues with embodiment in AR games that cannot be translated from traditional digital games or VR and require further research.

Subsequently, in Chapter 4, I zoom in on these issues and discuss various experiments concerning supernumerary bodies and body parts in AR. The first two pilot experiments concern the necessary forms of realism and mulitmodality to experience embodiment of a second virtual body and a third virtual hand. The third experiment combines these findings to show that embodiment of a augmented supernumerary third hand alongside the user's own hands is feasible. The fourth experiment studies the role of immersive tendency as a subjective measure of one's capability to become immersed in the experience of embodiment.

Finally, in Chapter 5, I examine the practical benefits of achieving embodiment in multimodal AR. Specifically, through an experiment I show that there is an indirect relation between a sense of body ownership over a virtual third hand and task performance through a sense of agency in a continuous action task.

1.2 Publications

This research has resulted in the following publications and master theses:

P1 Rosa, N. (2016, October). Player/Avatar body relations in multimodal augmented reality games. In Proceedings of the 18th ACM International Conference on Multimodal Interaction (pp. 550-553).

P2 Rosa, N., Werkhoven, P. & Hürst, W. (2016, November). (Re-)Examination of Multimodal Augmented Reality. In Proceedings of the 2016 Workshop on Multimodal Virtual and Augmented Reality (no. 2, pp. 1-5). ACM.

P3 Rosa, N., Hürst, W., Werkhoven, P. & Veltkamp, R. (2016, November). Visuotactile Integration for Depth Perception in Augmented Reality. In Proceedings of the 18th ACM International Conference on Multimodal Interaction (pp. 45-52). ACM.

P4 Rosa, N. E., Hürst, W. O., Veltkamp, R. C., & Werkhoven, P. J. (2017). Player-Avatar Link: Interdisciplinary Embodiment Perspectives. Encyclopedia of Computer Graphics and Games.

P5 Rosa, N., Van Bommel, J. P., Hürst, W., Nijboer, T., Veltkamp, R. C., & Werkhoven, P. (2019, March). Embodying an extra virtual body in augmented reality. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (pp. 1138-1139). IEEE.

P6 Rosa, N., Veltkamp, R. C., Hürst, W., Nijboer, T., Gilbers, C., & Werkhoven, P. (2019). The supernumerary hand illusion in augmented reality. ACM Transactions on Applied Perception (TAP), 16(2), 1-20.

P7 Rosa, N., Werkhoven, P., Hürst, W., & Veltkamp, R. C. (2020, March). A Model for Virtual Hand Ownership in Augmented Reality. In 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) (pp. 224-229). IEEE.

P8 Rosa, N., Veltkamp, R. C., Hürst, W., Brouwer, A.-M., Gijsbertse, K., Cocu, I., & Werkhoven, P. (2021) Embodiment and Performance in the Supernumerary Hand Illusion in Augmented Reality. Frontiers in Computer Science, 3: 694916. doi: 10.3389/fcomp.2021.694916

M1 Gilbers, C. (2017) The Sense of Embodiment in Augmented Reality: A Third Hand Illusion, Faculty of Science Theses, Utrecht University

M2 Van Bommel, J.G. (2017) Presence and Embodiment in Augmented Reality, Faculty of Science Theses, Utrecht University

Each chapter is based on one or multiple publications: Chapter 2 is based on P2, Chapter 3 on P4, Chapter 4 on P5, P6, P7, M1, and M2 and Chapter 5 on P8.

Chapter 2

A Theoretical Framework for Multimodal Augmented Reality

AR is commonly understood as a real environment that is augmented by virtual elements. Although there have been efforts in the past to concretely define AR and the elements it contains (Milgram and Kishino, 1994; Azuma et al., 2001), there are still cases of which the correct classification is open to debate. One such case, for example, is whether digitally captured and displayed elements are real or virtual. Another such case is when an application in which real occurs in one modality and the virtual in another should also be considered AR. Because of such debates there is no clear vision of what types of AR applications exist, or how fields have progressed since the introduction of AR, or what the future of AR research may hold. The purpose of this chapter is to disambiguate previous definitions of AR. In particular, I challenge the notion of 'real' and 'virtual' objects being the components of AR, the nature of augmentation and classify different forms of AR. The classification can be linked to application purposes and used as a tool to understand the minimal requirements of a system.

2.1 Real, Mediated, Virtual Stimuli

Although AR research as a whole has an immense corpus of literature, very little work has been dedicated to the exact differentiation between the terms real and virtual. Milgram and Kishino realized the need for such a distinction, and define real and virtual objects along three dimensions. Firstly, a real object exists objectively, whereas a virtual object exists in essence but not formally. Secondly, real objects can be viewed directly or indirectly, e.g. through a sensor-display mediator, whereas virtual objects can only be viewed indirectly. Lastly, a distinction is made between real and virtual images: real images have correct luminosity considering their surroundings, whereas virtual images have incorrect luminosity or are transparent.

The authors state that even this definition is not sufficient. In particular, it is noted that it is strange to consider both a remotely viewed video scene and one's own directly viewed hand as real, but do not further elaborate on this discrepancy. Some authors nowadays even address the former as virtual, possibly as a synonym for digital. This discrepancy is not further elaborated, but a similar discussion can be found in two other works, namely those of Mann (2002) and Müller (2015). Mann argues that a mediated view of the world (where mediation is not restricted to digitalization) is fundamentally different from a direct, undistorted view of the world. Therefore, any environment can not only be placed on a real-virtual continuum (in contrast to (Milgram and Kishino, 1994)), but also on a real-mediated continuum. Müller classifies information in procedural tasks in AR into five layers: the real world, the mediated world, virtual objects that are spatially referenced and of spatial nature, virtual objects that are spatially referenced but not of spatial nature, and virtual objects that do not have any connection to the physical world.

In essence there are two issues at hand here. Firstly, while AR is commonly described as consisting of real and virtual elements, in terms of perception and experience, mediated elements are something different and should be considered as a separate type. Secondly, Milgram and Kishino and Müller all discuss real, virtual and mediated in terms of objects in AR, while Mann discusses environments. Since today AR is not restricted to only augmentation of the real environment by virtual objects, but also allows for example modification of the real by virtual, a more appropriate element of discussion may be those elements that make up the perception of our environment, namely stimuli. Of course, this still allows the examination of complete objects, for example, in the visual sense, but also allows the examination of one's environment through other modalities. We can also now reformulate the existence dimension by Milgram and Kishino:

real stimuli originate from the physical environment, and virtual stimuli originate from a computer-generated model.

So what is a mediated stimulus in AR? Translating Mann's description of mediated to AR only, it is a digitally modified stimulus. The trivial case is using a sensor-display combination for illusory transparency. Examples of non-trivial modification, on the other hand, are: looking at your arm where the skin color has been changed digitally, or hearing someone's 'roboticized' voice, or sensing tactile sensations slightly intensified (Bayart and Kheddar (2006) define this as enhanced haptics). A more abstract example of modification would be sensory substitution, where a stimulus for a certain modality is presented as a stimulus for a different modality. Examples of sensory substitution are widespread; see Bach-y Rita and Kercel (2003) for an overview. The concept of sensory substitution can be extended to stimuli that humans cannot perceive (haptic enhancing in (Bayart and Kheddar, 2006)). Is it useful to point out that these forms of sensory substitution can also occur for virtual stimuli; an example can be found in *Luigi's Ghost Mansion* (2012) where the player senses when virtual ghosts are in close proximity through vibrations.

To limit the scope of 'modification' (i.e. the border between mediated and virtual), it is required that the modification is not random, and that there is a clear relationship between the modified and non-modified form. To illustrate, a point-cloud representation of a real person is considered mediated, whereas a 3D model of a generic human that is not derived from the real person in question is considered virtual, and of course the direct view of the person is real.

Putting together the above concepts, the following definitions can be derived:

Real the stimulus originates from the physical environment, and is perceived without any form of mediation, by its intended modality.

Mediated the stimulus originates from the physical environment, and is perceived through a digital sensor-display mediator, either by the intended modality (for which it may have been altered) or other modality(ies).

Virtual the stimulus originates from a computer-generated model, and is perceived through a digital display, either by the intended modality or other modality(ies).

There are a number of benefits to the differentiation between real, mediated and virtual. Firstly, it allows all applications for which the virtual means either 'digitally presented information' and/or 'computer-generated information' to still be considered AR. Secondly, sensory substitution stimuli, an important concept in multimodality which were formerly not considered real or virtual (with the exception being (Bayart and Kheddar, 2006)), can now be classified as an element in AR. More generally, any real stimulus that is digitally modified in some way can now be classified within AR. Lastly, it is likely that there is a difference regarding experience between the three types, for example, due to prior knowledge of one's environment, and applications can be designed accordingly. To illustrate, it is not strange to suggest that a user reacts more strongly to a threat to a mediated version of their arm than to a virtual arm. However, if the modification is large, or technology has progressed sufficiently, the user may not be able to make this distinction anymore. This utopian situation however is not expected anywhere in the near future, thus for the time being, it can be assumed this difference can be detected.

2.2 Basis and Augmentation

There are two influential works regarding the definition of AR: that of Milgram and Kishino (1994) and that of Azuma et al. (2001). Milgram and Kishino are widely known for their Reality-Virtuality Continuum, and define AR as all cases in which the display of an otherwise real environment is augmented by means of virtual objects. It is a subset of mixed reality (MR) on the continuum, which is in turn the merging of real and virtual worlds. Azuma et al. states that, indeed, a property for an AR system is that it combines real and virtual objects in a real environment. Furthermore he specifies that it must run interactively and in real-time, and it must register real and virtual objects with each other.

Both definitions agree in two aspects: the basis of AR must be a 'real' environment, and the augmentation is 'virtual'. According to the new stimuli-framework, there are two options for the basis, namely the basis is real or it is mediated. This distinction can already be found in the most common forms of visual AR technology: optical see-through HMDs provide a real environment as a basis, whereas video-based HMDs provide a mediated environment as a basis. The stimuli-framework then also allows augmentation by both mediated and virtual stimuli. To evaluate all possible combinations of basis-augmentation, I shall use the Holoportation technology by Microsoft (2016) as an example. This 3D capture technology allows precise, real-time reconstruction of humans, and in combination with AR displays can allow interaction with remote participants in 3D. The original demonstration shows a father wearing a see-through AR HMD who sees a digital representation of his daughter that is located in a different room, as if in the same room. Now, for each basis-augmentation combination, the following variations of this scenario hold:

- Real-Mediated - see-through AR where the interaction is with a real remote participant (as in the original Holoportation demonstration)
- Real-Virtual - see-through AR where the interaction is with a virtual agent
- Mediated-Mediated - video-based AR where the interaction is with a real remote participant
- Mediated-Virtual - video-based AR where the interaction is with a virtual agent
- Real-Mediated-Virtual - see-through AR where the interaction is with a real remote participant and a virtual agent.

In the last example, the mediated stimuli are part of the augmentation but generally they can be part of the basis, augmentation, or both. Although the combinations are sketched for one single application scenario, it should be clear that these variations hold for almost every other AR application, and that in other scenarios other combinations are possible.

What has been shown here is that while the previous definitions of AR still hold to some degree, the focus of AR has been shifted to a more generalized concept, namely basis-augmentation rather than real-virtual. In this way, AR is more clearly defined within MR, and augmented virtuality (AV) can be reintroduced (Milgram and Kishino, 1994). A requirement is that the basis and augmentation are related to the context of the application; simply adding unrelated virtual media does not validate the use of the term AR. In addition, complying with Azuma et al.'s other properties, I require that the AR system is interactive at real-time, and that real and virtual stimuli (rather than objects per se) are correctly registered with each other.

2.3 Multimodal Augmented Reality

Proposed concepts in visual AR literature are often said to be similarly applicable in multimodal situations. However, such conjectures are hardly ever verified and do not actually always hold. For example, imagine a navigation system where the virtual visual directions are overlaid onto the real visuals. Many would consider this system AR. Now, imagine the following alteration: the directions are presented as vibrations around the waist, similar to what is described by Van Erp et al. (2005). In the context of navigation, the visuals are the basis and the tactile directions are the augmentation. Although the alteration is slight, very few people would consider this version AR. Milgram and Kishino consider this case so different that they suggest using a new term altogether, namely 'hybrid reality'. This contrast is caused by the multimodal factor.

A small group of works have been dedicated to the closer examination of multimodal AR, and in particular, the larger number of variations it allows compared to unimodal AR. Kalawsky et al. (2000) propose a framework based on a functional decomposition of AR. The purpose is to allow complete sensory description of the user's capabilities to compare systems. Although the taxonomy can be used to describe a specific multimodal AR application, it is in fact not specific for multimodality nor for AR. It is a decomposition based on any generic human-computer interface, and there is no mention of 'combining real and virtual' or 'augmentation' for that matter. Lindeman and Noma (2007) provide a classification system for AR based on where the real-virtual mixing occurs, which can be used for each sense individually. They give examples of technologies for various mixing locations and describe important implications of mixing location, such as the fact that mixing location and technology choices made for one modality can constrain the options for the remaining modalities. However, the classification system is restricted to those cases where the virtual must be as realistic as possible and does not consider, for example, more abstract forms of sensory substitution. Jeon and Choi define a visuo-haptic MR taxonomy, which consists of two orthogonal reality-virtuality continua (from (Milgram and Kishino, 1994)), one for visuals and one for haptics. Unfortunately, it extends poorly to all modalities, since the complexity grows exponentially with each added modality.

What these works so far do not illustrate, with (Jeon and Choi, 2009) being the exception, is that multimodal AR can indeed mean that the basis and augmentation are within the same multiple modalities, but also that they are spread across modalities. Previous visual AR research has been more concerned with correct implementation of perceiving in high quality, but when the goal of an AR application is to recreate a real scenario, then congruence across modalities is also of importance. However, this goal is not necessarily generalizable to all AR applications.

2.4 Inter-, Intra-, Crossmodality

In the presented classification system, the different categories depend on the degree of spreadness of basis and augmentation across modalities. The classification is therefore not concerned with the quality of the stimuli, the number of relevant utilized modalities, or the quality of blending of the basis and augmentation. All that is required is that the basis and augmentation are linked by the purpose

AR Case	**Vision**		**Audio**		**Haptics**		**AR Example**
	Basis	**Aug.**	**Basis**	**Aug.**	**Basis**	**Aug.**	
Intramodal	✓	✓	✓	✓			Holoportation (2016)
Intermodal type 1	✓	✓			✓		Soft AR (Punpongsanon et al., 2015)
Intermodal type 2	✓	✓				✓	Luigi's Ghost Mansion (2012) for AR
Intermodal type 3	✓	✓		✓	✓		Soccer Scenery
Crossmodal	✓					✓	Visual Navigation with Tactile Belt (Van Erp et al., 2005)

Table 2.1: The classification of different types of multimodal AR, illustrated by examples using vision, audio and/or haptics.

of the application (as before). Table 2.1 illustrates the classification and gives examples of AR applications for each category, which will each be further elaborated in the following. These examples only regard vision, audio, and/or haptics, but the classification works for all modalities, including those outside of the traditional five. The basis-augmentation model can moreover be applied to both AR and AV, the two subsets of MR. Therefore, a classification system based on this model can be applied to MR in general. In the following, I shall describe the system in the context of AR for simplicity.

To start, there are at least two cases: the basis and augmentation are within the same modalities (e.g. basis and augmentation in both vision and audition) or they are not (e.g. visual basis and auditory augmentation). These two cases are called intramodal and crossmodal AR, respectively. Intramodality is the generalization of unimodal AR to multimodal AR, where for each modality related to the application there is a real/mediated basis and a mediated/virtual augmentation. Generally, intramodality can be useful for scenarios where the purpose is to have the final augmented scene as close to a realistic one as possible. Returning to the Holoportation example, the purpose of AR in this scenario is to create the sensation that the daughter is actually present alongside the father, and therefore it is important to include both basis and augmentation in those modalities that matter. In the demonstration video, the father sees and hears this mediated daughter, which creates the impression she is there. Crossmodality requires that the modalities of the basis and augmentation are mutually exclusive, that is, there is no overlap in modalities. An example of this is the visuotactile navigation scenario described earlier. Of course, the user is still capable of perceiving other real stimuli, such as real tactile sensations, however these stimuli are not crucial to the application, so they are not considered as a basis modality. In this example, the purpose of crossmodality could be to ensure high performance in a situation where it would be otherwise disadvantageous to impair the visual basis with visual augmentation, and thus moving certain information to a different sense.

The last type of AR is the middle ground between ‘not any’ and ‘complete’ spreadness across modalities. Intermodality indicates that there is intramodality, and either a basis in one or more different modalities, an augmentation in one or more different modalities, or crossmodality. An example of the first subtype is SoftAR, where a user can see and touch a real object, and virtual indent marks are projected onto the object such that the user experiences it as softer than it really is (Punpongsanon et al., 2015). An example of the second type would be an immersive AR variation of Luigi’s Ghost Mansion, where a player has to locate the virtual ghosts, of which the presence is felt through the tactile sense, and they appear and are defeated when the player looks straight at them. Lastly, an example of the third type would be a real game of soccer but the surroundings of the field are augmented to represent a virtual stadium where virtual fans are chanting.

To summarize, the classification characterizes the following types of AR, from least to most spreadness across modalities:

Intramodality all bases and augmentations are within the same modalities

Intermodality intramodality in at least one modality and:

- **type 1** a basis in one or multiple different modalities
- **type 2** an augmentation in one or multiple different modalities
- **type 3** crossmodality in other modalities

Crossmodality the modalities of the bases and augmentations are mutually exclusive.

This classification is useful to gain insight on different areas of research being conducted in multimodal AR. It is generalizable to all modalities, and can be restricted when necessary to a specific group of modalities as done in Table 2.1 and by Jeon and Choi (2009). Another benefit is that it can be linked to application purposes and used as a tool to understand the minimal requirements of the system. For example, when an AR scene is required that mimics the real world as close as possible, intramodality is likely desired. On the other hand, when it is necessary to not overfill or modify a certain basis modality, crossmodality offers valuable implementation options. Intermodality may be used when a combination of these two goals is necessary. I emphasize that these are rough generalizations, and recognize that exceptions exist.

The goal of this classification is to complement earlier classifications, creating a more overarching view of what multimodality in AR, and more generally MR, can be. For example, the taxonomy described by Jeon and Choi (2009) is very similar in nature, but a few key differences are noticeable. Firstly, the cases that they describe as MR are only cases where there is intra- or intermodality (types 1 and 2). The currently proposed classification complements this by considering their ‘rV-vH’ and ‘vV-rH’ as crossmodal MR. Secondly, the third subtype of intermodality which concerns three or more modalities can also be classified in MR.

2.5 Conclusion

In this chapter I redefined aspects of typically unimodal AR such that they can be used for multimodal AR. In particular, it is argued that the elements of AR are not only real and virtual stimuli, but also mediated stimuli. Moreover, the core of AR is not simply the combination of these stimuli, but more generally a real or mediated basis that is augmented by mediated or virtual elements. Viewing AR in such way allows us to reconsider the notion of AV, and in turn the notion of MR. I presented a classification system for multimodal AR based on the basis-augmentation model.

In the next chapter, I describe what this new system for defining and classifying multimodal AR could mean for body-related experiences in AR. To do this, I first discuss the relation between players and their in-game body representation, the avatar, in conventional digital games. Then, I discuss one type of relation in particular, embodiment, and how the current knowledge on embodiment may not exactly translate to AR games, and the difficulties that must be overcome to accomplish embodiment in AR. In Chapters 4 and 5, I show through various experiments the existence and function of embodiment in AR.

Chapter 3

The Player-Avatar Link: Interdisciplinary Perspectives

3.1 Introduction

In digital games, players are typically represented in-game by an avatar. While it is widely acknowledged that some type of link exists between the player and their avatar, there is still much disagreement on what the nature of this link actually is. This chapter discusses the different perspectives of the player-avatar link. Specifically, I first examine phenomenological accounts of player embodiment and views beyond embodiment in games from a Humanities background, and then investigate the sense of embodiment in virtual media from a Natural and Social Sciences background.

3.2 Terminology

A few concise definitions are provided as a starting point for the main discussion. These definitions are simplified versions of what they may represent within different disciplines.

Avatar The object over which the player can assert control.

Character The figure that exists in the meaningful game world, including any form of backstory and personality.

Embodiment The process of adjusting one's internal body representation to the current circumstances.

Player embodiment The experience that the avatar has changed the internal body representation and phenomenal body of the player.

Presence The experience of being present to something.

Spatial presence The experience of being present in a certain environment.

Immersion The technological quality with respect to sensorial information.

In the following the term 'game figure' is used to refer to either avatar and/or character, when it is not specified by the author(s) of the cited article.

3.3 Perspectives within the Humanities

In the Humanities there are several authors who acknowledge that the player-avatar link is exactly embodiment, in the form of extension of the player's body. In a number of his works, Wilhelmsson (2001, 2006, 2008) has argued that many aspects of a game experience are consequences of a Game Ego manifestation. The Game Ego is a bodily based function that enacts a point of being within the game environment through a tactile motor/kinesthetic link. This means that the player's sensory system is extended to the game environment, and the Game Ego becomes another body and/or an extension of the body. The consequences include identifying with the manifestation, and more generally allowing the player to experience the game's narrative elements, and as a result evoking emotional responses and the experience of presence. Klevjer (2006, 2012) states that there is a paradoxical prosthetic relationship between player and avatar. By applying Merleau-Ponty's (1945) philosophy on the body's duality, it is argued that video games allow the player to relocate their intention for actions (as a subject) into screen space, while at the same time a proxy (as an object) exists in the game. Crick (2010) also draws upon Merleau-Ponty's reasoning to explain how the game can be perceived as another physical world during play. That is, the player exists in two worlds, and operates both on and in the game. It is emphasized that the video game experience must be of embodied perception, since players still require their body (and the combination of all senses) to perceive it, and it in turn affects their bodily state. The controller used by the player allows agency in the game world and can become an extension of their body through habit; the avatar's movement is incorporated within the player's body schema, and becomes an extension of the bodily basis of consciousness. Lastly, Vella (2013) argues that, indeed, the player may embody the avatar and achieve a subject position in the game world, but as a character the game figure still has an autonomous identity that can be acted out. Therefore the act of avatar-play constitutes the enactment (or performance) of a character.

These works so far demonstrate an emphasis towards the roles of perception and cognition in embodiment: their combined impact on our experience of the real world is considered in the experience of game worlds. Farrow and Iacovides (2014) claim that exactly such reasoning is just one step too far, and describe three inconsistencies with respect to human embodiment aspects laid out by Merleau-Ponty. Firstly, on a physical level bodies within game worlds cannot conform to real world duality (e.g. tactile and pain), and we do not relate to bodies in virtual worlds in the same way that we do in the real world. Secondly, on an intentional level a player experiences the game as convincing when there is a sense of non-mediation, which is something only 'invisible' game control systems can achieve. Lastly, on a worldliness level it is possible that game worlds only become meaningful through play with other humans. Together, these inconsistencies leads to a limit to the degree of digital embodiment.

There are also authors that explain that embodiment in games is not 'simply' a process of perception. For example, Newman (2002) has argued that "video games are not interactive, or even ergodic", since they do not consist of continuous play and even have integral parts that are non-ergodic. Here, ergodicity refers to the definition by Aarseth (1997): a user must use (active) effort in order to experience the medium. Moreover, there is a level of ergodicity in non-controlling players (active spectators), which indicates that feelings of immersion, engagement and being-in-the-game are separated from an interface-level control loop. During play, the degree to which the player embodies the game figure is not dependent on representation, since it is merely seen as a set of capabilities: it is equipment to be used by the player.

A last group of authors agree that the player-avatar link takes on more forms than just embodiment, and that this is very related to different types of play. Linderoth (2005) demonstrated how children frame the game figures during certain moments of gameplay, which leads to three different functions of game figures: they can become roles for socio-dramatic interaction, tools as extensions of the player agency, and props for self-presentation in the presence of others. Similarly, Bayliss (2007) argues there are three positions of game figure play: playing through (the avatar is equipment), playing as (sometimes character, sometimes avatar), and playing with the game figure (play with the game rather than play the game). The game figure embodies the intentions of the player as their avatar, and its limitations with respect to functionality simultaneously constitute it as an embodied character. It is highlighted that any sense of being-in-the-game-world relies on attitude of the player with respect to the three positions of play, and not the video game itself or the technological platform. Lastly, Banks (2015) found that players' motivation and attitude towards play went hand in hand with the social role the game figure fulfilled in the relation with the player. From unsocial to social, the relations were game figure as object, as me, as symbiote, and as social other.

To summarize, there is a substantial group of authors that argues for embodiment of the player as a result of a perceptual link: the player's body is extended into the game through the avatar. However, others argue that real world embodiment assumptions cannot simply be translated to game world assumptions, and that the function of the game figure depends on the type of gameplay, which in turn depends on the player's attitude towards the game.

3.4 Perspectives within the Natural and Social Sciences

In the Natural Sciences there is little attention to the player-avatar link in games specifically, however there are many works that discuss how the body schema can be changed by virtual bodies in, for example, VR. This corpus is a result of empirical studies in cognitive neuroscience on the more general experience of being connected to a body. Although there are many concepts that are part of this experience, there are three in particular that have gained a great deal of attention, that together form the sense of embodiment: body ownership (the sensation of owning a body), agency (the sensation of controlling a body), and self-location (the sensation that the locations of you and your body coincide in space) (Kilteni et al., 2012a; Longo et al., 2008). Here, spatial presence differs from self-location since the former concerns the relation between the self and the body, and the latter the self and the environment. A classic experiment to assess ownership over a limb is the RHI (Botvinick and Cohen, 1998). In this experiment the rubber hand is stroked either synchronously or asynchronously with the real hand, which is out of sight. Synchronous feedback evokes a sense of ownership over the rubber hand, while asynchronous feedback diminishes it. With similar setups for not just limbs, but also entire bodies (i.e. the Full Body Illusion), various studies have assessed the importance of seemingly relevant factors to the illusion, such as synchronous visuotactile and visuomotor feedback (Tsakiris et al., 2006), viewing perspective (Petkova and Ehrsson, 2008) and congruent body alignment and connectivity (Perez-Marcos et al., 2012). It has also become apparent what the roles of agency and self-location are in this regard (Kalckert and Ehrsson, 2012; Maselli and Slater, 2014; Tsakiris et al., 2006).

Using VR it became possible to inspect real-world factors that could not, or with much difficulty, be studied otherwise. Lugrin et al. (2015) examined the full body during a full body touch-the-target task in VR, where ownership over humanoid, robot, and block avatars were compared. They found that there was no difference in ownership levels between the different avatars (the actual levels are not provided), however the humanoid avatar caused the participants to experience having two bodies to a higher degree than the other non-humanoid avatars. The authors believe that this effect was related to the Uncanny Valley Effect (Mori et al., 2012). Besides full body illusions, the extension and addition of body parts has also been examined. For example, Kilteni et al. (2012b) examined how the degree of ownership over a virtual arm depended on elongation of the arm, and found that with visuotactile feedback the breaking point was four times arm's length. Regarding supernumerary limbs, Steptoe et al. (2013) showed that participants could experience ownership and agency over a humanoid avatar with a long tail, while also performing better in a full body touch-the-target task when the tail could be controlled by hip movement than when it moved at random. In a similar study, Stevenson Won et al. (2015) compared performance in an arm-based touch-the-target task between a humanoid avatar and one with an extra arm protruding from the chest, which could be controlled by wrist rotations. For targets outside of the normal arm's reach, but in reach of the additional arm, participants performed significantly better with the extended body than with the normal body. Interestingly, the measured levels of presence were low overall and did not differ between conditions. The authors explain that the low results could have occurred because participants were so involved with controlling the avatar in order to complete the task, making them less aware of the virtual surroundings. This is in contrast to many works in Natural Sciences that agree that there is in fact a positive relation between presence and task performance (an elaborate discussion can be found in (Nash et al., 2000)).

Although measuring embodiment in games directly (i.e. through ownership, agency, and self-location) is uncommon, there have been many works that measured the effect of presence and/or immersion in games on a variety of concepts. For example, in a series of studies, Weibel and Wissmath (2011) studied possible influences and effects of spatial presence and flow (i.e. being immersed in what your are doing) in a variety of games using both factor analysis and path analysis on questionnaire results. Firstly, they found that spatial presence and flow are separate constructs, and that flow in turn consists of two subcomponents: absorption into the experience and smoothness of the experience. Secondly, they found that a participant's motivation generally influences flow, whereas a participant's immersive tendency generally influence presence. Lastly, flow directly influenced enjoyment and performance of each game, whereas presence only did this indirectly through flow.

Besides subjective qualities of gameplay, objective qualities have also been examined. For example, Hou et al. (2012) studied players' immersive tendency as a precondition of presence experience, and screen size as a media form variable. By analysis of questionnaire results they found that a larger screen had a positive effect on game figure evaluation, player mood, and both spatial and self-presence (i.e. when a player experiences that the avatar is him-/herself), but no difference in enjoyment. Also, immersive tendency moderated the effect of screen size on presence, but not of the other game evaluation aspects. In particular, enjoyment was not affected, in contrast to the study by Weibel and Wissmath.

To summarize, although there is little research on the player-avatar link itself, there is an immense corpus of literature regarding the general experience of being connected to a body in reality and VR that prominently argue from a perception background: multimodal feedback can cause participant' s to experience other (virtual) bodies as their own. There have also been studies that empirically show that a player's immersive tendency are an important determining factor of the experience of spatial presence in games.

3.5 Body and Environment

Although the different disciplines study the same problem very differently, there are a few interesting similarities between the two, specifically regarding their interpretation of the problem. For example, both disciplines agree that during gameplay the player's internal body representation changes. In Humanities some authors draw upon the phenomenology of natural perception, and argue that a part of the player must be represented in the game. In Natural Sciences the reasoning is that if extra bodies (or body parts) are presented correctly, then they can be accepted by the player as belonging to them. Both sides argue for an extension of the player's body, and many results support this claim: players frequently refer to the avatar as 'I', react emotionally and physiologically to avatar events as if it is their real body (Armel and Ramachandran, 2003), while still experiencing ownership over their actual real bodies (Guterstam et al., 2011).

Another similarity is the important role of the experience of the game environment, in particular what the determining factor is of this experience. In computer science it was until recently quite acceptable to regard presence as a consequence of a system's immersive and interactive capacities and of nothing else. This is often a crude operationalization of Steuer's (1992) model on presence in VR (Steuer, 1992); the statement concerning varying individual experiences of presence due to, for example, differences in the content, is frequently ignored. Currently, a progressively growing number of studies are demonstrating the importance of the player's immersive tendency to the experience of the environment. In game studies, system-versus-narrative discussions are nothing new: is the experience of spatial presence (or being-in-the-game-world) a result of the vividness and interactivity of the system, or is this a result of the environment becoming meaningful through narrative? There are numerous authors that are inclined towards the latter (Bayliss, 2007; Farrow and Iacovides, 2014), although undoubtedly the answer is partially both. A possible reason for the exclusion of this view in computer science is one of pragmatic nature: there is not yet an existing measure for 'the degree of narrative'. This is absolutely logical when looking at the complexity of experiencing narrative; it can differ per person, per experience of gameplay, time of day, and so on.

For game research in both disciplines, the experiences of self, body and environment often go hand in hand. In many cases, it seems to roughly come down to two aspects: the mediation and immersive tendency. That is, the only way a person could experience a mediated or virtual situation as real is if the input is *realistic* or convincing enough, and if they allow themselves to *pretend* the experience is not mediated or virtual. It is not apparent what the interdependence is between these two aspects, nor whether they are of equal importance: acceptance in the second aspect allows experience of the first on the one hand, but it may very well be that if the first occurs perfectly then the second happens automatically (as suggested in (Wilhelmsson, 2008)).

3.6 Augmented Reality Games

There are two crucial differences between the nature of the traditional digital games and VR environments described above and AR games. Firstly, the game world is now a merged world in which the both player and avatar are situated. Secondly, the game world can take on different combinations of basis and augmentation. In the following I describe how these differences change the idea of embodiment as a player-avatar link in AR games.

In traditional digital games the player and avatar were separated by a non-immersive screen and controller. Many games showed an in-game game figure seen from a 3PP, where the whole game figure body was visible, while others also showed an implicit avatar seen from a 1PP, possibly with a visible body when looking down. VR games make use of an HMD's immersive capacities and typically use a 1PP, where the player's body is replaced by an avatar body. Both of these avatar representations, non-overlapping and fully overlapping, can be translated to AR. In the first case, the player sees a game figure in the merged game world, separated from their own body. An example of this can be found in the game Young Conker for the Microsoft Hololens (2016), where the player controls the direction in which Conker moves using head orientation (i.e. the center of their visual field). In the second case, there is an implicit avatar since the player takes on the role of the protagonist. An example of this can be found in another Micosoft Hololens game, Fragments (2016), where the player represents a detective that has to solve a crime using memories, and the crime scene is integrated into the player's surroundings. Besides these two existing forms, there is another that is unique for AR, namely a partially-overlapping avatar. Currently no example of such an avatar exists in commercial AR games, but one possible form would be the player's real body being augmented with virtual limbs. This resembles 1PP digital games where only the arms of the game figure are visible to the player, as in the famous computer game Doom (1993). One crucial difference is that in such computer games the game is not necessarily implying that the on-screen arms belong to the player, whereas this is surely possible in AR, by for example visually connecting the virtual limbs to the real body.

For each case, I infer whether the above described views from all disciplines still hold or can be adjusted for AR games. For the non-overlapping avatar case, the player and avatar's bodies may be less 'far apart' since they are now both in the game world. Although Crick's reasoning would suggest this is still a case of embodied perception, both Wilhelmsson and Klevjer's view rely on a game world that is separated from the player. It is difficult to say whether being in the same world would imply a stronger or weaker form of embodiment. On the one hand, you are closer to the game figure than before which could evoke a stronger sense of embodiment. On the other hand, although you as a player are still fully in control, seeing the game figure in the same environment as you are may make it feel more like a being on its own, or even less like a living thing and more like a tool. Indeed, studies on full body illusions in VR would suggest that a 1PP is crucial to the experience of body ownership, and is completely diminished in a

3PP, however these do not typically simultaneously show the player's real body and fake body. An experiment by Petkova and Ehrsson (2008) showed that when shaking another person's hand, but seeing from that other person's perspective, there was more ownership over the other person's body than their actual body, but the comments would suggest there was disownership of the actual body. In other words, the 'other' body as seen from afar is indeed not embodied.

In the fully overlapping case, there is little reason to believe the situation is anything other than embodiment, and body transfer illusions in VR using a 1PP support this notion. The partially overlapping avatar is not as straightforwardly translated, but would also suggest an experience similar to embodiment. SHIs in reality have shown that embodiment of a third arm is possible without disembodiment of the real hand (Guterstam et al., 2011), but there are many more variations to consider that may not all lead to the same degree of experienced embodiment. For example, in the case of a virtual limb, an arm can be connected to the user from the chest, or it could just be a floating hand with invisible arm but still clearly is represented as part of the player's body.

Now, besides differences in avatar representation in AR, the nature of the merged environment must also be considered. Both Hololens examples of AR games given above, Young Conker and Fragments, use a real basis (since it is an optical see-through HMD) and a virtual augmentation. One could argue that such divergence between basis and augmentation may negatively affect the overall perceived 'realism', which in this case refers to how well the basis and augmentation seem to create a seamless whole. The traditional approach to this problem is to perfect the virtual as to match its real world counterpart: make the virtual look/sound/feel as 'real' as possible. Unfortunately, technologically speaking, this is not yet possible. There is, however, a different approach: narrow the gap between basis and augmentation. This has been studied in the visual sense, where one study in particular stylized both basis and augmentation to look cartoon-like in order to blur their boundary, but was not further tested on users (Fischer et al., 2005). Another study measured that enhancing vertical and horizontal edges of an entire merged scene, as if to add cartoon-like outlines, indeed lead to chance-level discernability judgments between basis and augmentation, that is, participants could not tell whether an object was real or virtual (Steptoe et al., 2014). In terms of embodiment, one could reason that if the player of an AR game cannot tell the difference between basis body parts or augmentation body parts, then there may be a greater tendency to experience embodiment. In the example of a third arm, seeing the real body with an optical see-through HMD with an abstract virtual arm model would be an instance of a wide gap between basis and augmentation. Examples of closer basis-augmentation combinations could be:

- real-mediated: seeing the real body with an optical see-through HMD and a digital projection of the user's actual arm presented as third arm;
- mediated-virtual: seeing a stylized mediated body by means of a video see-through HMD with stylized virtual third hand model;
- mediated-mediated seeing a stylized mediated body by means of a video see-through HMD and a digital stylized projection of the user's actual arm presented as third arm.

To summarize, in terms of embodiment and avatar representation, for non-overlapping and partially overlapping avatars it is unclear whether evidence of embodiment as in traditional digital games and VR games can be translated to AR games. Moreover, it is possible that by bringing the basis and augmentation closer together the tendency to experience the sense of embodiment may be greater.

3.7 Conclusion

This chapter has provided an interdisciplinary discussion of the player-avatar link. For Humanities there is no consensus on the importance of perception over player attitude and narrative. In Natural and Social Sciences the experience of changing the body schema through perception is prominent, however there is recently also noticeable focus on personal characteristics that might underlie the experience in the first place. Both disciplines are inclined to argue about the experiences of both body and environment, and end up at the same important question: what are the roles of mediation and immersive tendency in these experiences? For AR games and embodiment specifically, this includes the choice of basis and augmentation, and the unexplored areas in terms of avatar representation. These issues are the focus of Chapter 4.

Chapter 4

Supernumerary Embodiment in Augmented Reality

As discussed in the previous chapter, there are a few crucial considerations when creating embodiment experiences: the avatar representation, the choice of basis and augmentation, and the role of immersive tendency. In this chapter, I describe four experiments on the embodiment of supernumerary hands and bodies in AR.

I first discuss SHIs in reality, supernumerary body illusions in digital games, and hand illusions in AR specifically. Then, the first two pilot experiments are described, where each considers a single form of avatar representation (overlapping in experiment 1 and non-overlapping in experiment 2) and the choice of basis and augmentation (real-virtual in experiment 1 and mediated-virtual in experiment 2). These aspects were split across experiments since the goal was not to compare them but to inspect them as individual experiences. Moreover, in experiment 1 multiple forms of multimodal stimulation were examined, since it was unclear from related work which would be required to illicit an embodiment illusion in AR specifically. The approach was to use the resulting required stimulation in experiment 2, but due to problems arising from the hand appearance in particular, this was further examined there as well. Then, in the third experiment, I combined optimal settings based on these pilot experiments in order to show that it is possible to elicit a sense of embodiment over a disconnected supernumerary mediated hand simultaneously visible with the user's mediated body. In a final follow-up experiment, I show that the variation in subjective experience of embodiment from the third experiment could be explained by the participants' self-rated immersive tendency.

4.1 Supernumerary Hand Illusions in Reality

Armel and Ramachandran (2003) implicitly study the SHI when examining the classic RHI and a so-called 'table-illusion' where it was tested whether participants could feel ownership over a tabletop (Armel and Ramachandran, 2003). Using the visual presence of the real hand as a control condition, the authors reported significant differences in both subjective illusion rating and objective skin conductance response (SCR) after a threat to the fake hand that together indicated that the illusion can only occur when the real hand is hidden and not when it is visible, and that it only occurs subjectively for a table.

Other studies, in contrast, have found that participants can still feel ownership over a fake limb when the real hand is visibly present. Schaefer et al. (2009) asked participants to look at both of their real hands and a rubber left hand, and stimulated the pinky and thumb of the real left hand by synchronous touching; all other hands were unstimulated, including the fake hand. There were three conditions: connected rubber hand, disconnected rubber hand, and no rubber hand. The authors measured subjective experience, and the angular distance between activated areas in the somatosensory cortex corresponding to the pinky and thumb. They found that when the rubber hand seemed connected to the body, the angular difference was smaller than when they seemed disconnected and when there was no rubber hand at all. These results are supported by similar significant differences in the illusion ratings, although the overall ratings were quite low. This shift in cortical representation of the thumb indicates that the somatosensory homunculus reflects the perceived shape of the body rather than the physical aspects of peripheral stimulation as previously thought, and in turn that it is in fact the shape of the body that changes when multimodally presented with extra body parts. Another study by Guterstam et al. (2011) demonstrated that ownership can indeed be induced over a rubber right hand when the real right hand is visible, without disownership of the real hand, but that this coincided with less ownership of the rubber hand compared to an RHI with hidden real right hand. By means of questionnaires and SCR, they showed that by using synchronous stroking a higher level of ownership was achieved over a rubber arm compared to asynchronous stroking, and that ownership did not occur for other limbs or arms of incorrect rotated position. Lastly, Chen et al. (2018) showed that when participants sat across from the experimenter and saw four hands, with the experimenter's hands in 1PP and their own in third, which they tapped on the table surface synchronously while being synchronously stroked, they subjectively had the experience of owning four hands. This illusion was not reflected in the objective SCRs, and did not occur for only brushing or only tapping, or when the participant saw their own hands in 1PP.

A last category of SHI studies has used two fake hands while the real hand is hidden. Ehrsson (2009) used two rubber hands that were placed above and slightly to the left and right of the real right hand, with the left hand also in view. The results showed that there was a significant difference in SCRs between synchronous and asynchronous stroking, confirming that healthy individuals are capable of feeling and seeing supernumerary limbs. Newport et al. (2010) explored how the SHI affects both the body schema (body representation for action) and body image (body representation for perception). The authors used two video displays of the real left hand, which is in turn out of sight, and synchronously stroke either the rightmost or leftmost hand (other asynchronously), or both synchronously with the real hand. The effects on the body image were measured by means of a questionnaire, and the effects on the body schema using a pointing task. The questionnaire showed that participants experienced ownership over the hand/hands that was/were stroked synchronously, while the pointing data only supported single hand ownership. The authors concluded that both body image and body schema can accommodate a fake limb, but only the body image supports multiple fake limbs. Finally, Folegatti et al. (2012) investigated multiple fake limb ownership in the context of the body's spatial constraints. Using two rubber right hands placed to the left and even more to the left of the real right hand (both real hands out of sight), they show that when one rubber hand is stimulated synchronously and the other remains unstimulated, ownership can occur over the synchronously stimulated one, regardless of how far it was from the real hand, according to questionnaire ratings and proprioceptive drift measures. Furthermore, when both hands are stimulated, only the closest rubber hand can be owned. The authors conclude that we should consider spaces that 'belong to the body', 'can belong to the body', and 'can affect the body' in a new conceptual framework in the context of ownership.

4.2 Supernumerary Body Illusions in Digital Games

Full body illusions do not typically focus on supernumerary bodies. Instead, the focus is on whether a different full body that visually replaces the own body can be embodied, and which factors are necessary for this illusion. For example, Slater et al. (2010) showed that 1PP and visuotactile synchrony were important factors to generate a full body illusion, where the former dominated the latter; embodiment also occurred for both synchronous and asynchronous visuomotor synchrony of the head. In another study, it was shown that a 3PP of the mediated own body and a 1PP of an invisible body being synchronously stroked led to disownership of the real body (Guterstam and Ehrsson, 2012). In a body swap illusion, Petkova and Ehrsson (2008) showed that when another person's body is viewed in a 1PP, shaking hands with one's own body viewed from a 3PP, then the other person's body is owned, but the actual own body is not.

Game studies where players control in-game avatars typically do not measure embodiment with an RHI paradigm. Instead the focus is aspects such as presence, flow, and enjoyment. A term that is seen often in relation to presence, which is focused on the experience of the virtual environment, is the concept of self-presence, which is focused on the experience of the virtual self (Biocca, 1999; Lee, 2004). This is very similar to the sense of embodiment, but differs in its scope as can be seen from its arrangement of components. Proto self-presence, the sense of physical being, resembles body ownership, but the other components core and autobiographical self-presence, relating to emotions and identity, respectively, do not coincide with agency and self-location at all (Ratan, 2012). In a study by Ratan and Dawson (2016), participants were asked to play Wii Sports Resort Swordplay, where a humanoid Wii avatar, a Mii, was controlled by the participant's motion as registered by the Wii remote. After play, the players watched the Mii get beaten without being able to intervene or control the avatar. The results showed, amongst other findings, that participants who experienced more proto self-presence during play also experienced less avatar self-relevance (i.e. perception of the avatar being relevant to the self) when the avatar was threatened after play, and conversely, those who experienced less proto self-presence experienced more avatar self-relevance. The authors explain that the contrast between movement congruence during play and movement incongruence during non-play was so large for those who experienced more proto self-presence, that any avatar self-relevance was diminished; this contrast was not so large for those who already experienced less proto self-presence, thus self-relevance was better retained. Although the focus of this study is not self-presence, it indirectly shows that self-presence can indeed be experienced during gameplay where the avatar body is separated but also simultaneously visible to the player, if there is congruent motor control. Importantly, it also emphasizes that there is great variance in experienced self-presence among players. Other studies have shown that different levels of self-presence are related to different levels of parasocial interaction with in-game avatars (Jin and Park, 2009), predicting exercise accomplishment in exergames (Song et al., 2014), and offline health and appearance (Behm-Morawitz, 2013).

4.3 Rubber Hand Illusions in Augmented Reality

Although many RHI studies take place in reality or VR, comparably there seems to be less interest in investigating the AR case, or less means to due so by technological constraints. One of the few studies is the one performed by IJsselsteijn et al. (2006), where the RHI is performed in three forms of media, all with synchronous brush stroking: reality, 'VR' (2D tabletop projections of rubber left hand and brush), and 'MR' (2D tabletop projection of rubber left hand, real brush). We remark that despite the names, both the 'VR' and 'MR' conditions could both be considered variations in AR, according to the definitions in Chapter 2: there is a real basis (table and right arm) and a mediated augmentation (left hand). One of the main goals of the study was to explore how well the RHI can be reproduced in various media. Questionnaire and proprioceptive drift results showed that ownership occurred in all cases, albeit weaker in the mediated conditions, likely due to the incongruent hand shape (i.e. flat). Moreover, according to the questionnaire, ownership in the VR case was stronger than in the MR case. The authors explain that this could be due to inconsistencies in the MR case as a result of the not so seamless integration of real brush and mediated fake hand, which led to incongruent stroking texture.

Suzuki et al. (2013) studied the interaction of exteroceptive with interoceptive processes in a 'cardiac RHI' using an AR setting. The authors captured depth and color information of the real left hand, which was out of view. With this, participants could see their mediated hand in a mediated environment. There were six conditions: synchronous/asynchronous cardiac pulsing without movement, synchronous/asynchronous cardiac pulsing with movement, and synchronous/asynchronous visuotactile stroking. The results showed that in both synchronous cardiac feedback cases, the participants subjectively reported a sense of ownership. Moreover, when participants were able to move the mediated hand, ownership also occurred in the asynchronous cardiac feedback case. Still, the synchronous visuotactile case without any cardiac feedback resulted in the highest reported ownership, even though the mediated hand became 'broken' (i.e. discontinuous) when the mediated brush appeared. The authors conclude that multisensory integration of exteroceptive and interoceptive signals can modulate body ownership, but suspect that movement signals dominate the influence of the interoceptive signals. Importantly, they suggest that a priori willingness to accept the mediated pulsing hand could have played a role, as did interoceptive sensitivity. In a more recent study, Feuchtner and Müller (2017) investigated how users could control far away mediated objects using a virtual arm in AR. They examined four types of visualizations: a disconnected virtual hand without seeing their own mediated hand (removed in the style of diminished reality), disconnected virtual abstract hand pointer without own mediated hand, connected virtual hand without own mediated hand, and a connected virtual hand with own mediated hand. The latter condition could be considered a supernumerary hand illusion setup. Their results showed that ownership only occurred in the condition where the virtual arm was shown without the own mediated hand; other conditions, including the SHI, showed negative ownership results. They concluded that sufficient hand realism and connectedness preserve ownership in AR, whereas seeing the own mediated hand disrupts the illusion.

Lastly, some studies have investigated embodiment scenarios in AV. In one particular study, Jung et al. (2018) investigated the influence of virtual body representation on object size estimation using conditions comparing an AV scenario with a personalized mediated hand image to a VR scenario with a generic virtual hand model. They showed that the personalized hand image increased both subjective ownership of the hand and spatial presence compared to a generic virtual hand model, and furthermore supported participants in correctly estimating the size of a virtual object in the proximity of their hand.

4.4 Purpose

A few things are apparent from this overview of studies. Firstly, since the results from SHIs are very mixed, it is still unclear whether one can embody supernumerary hands. Secondly, while there is at least some evidence that supernumerary hands can be embodied while the real own hands are visible, there is even less evidence for supernumerary bodies. Lastly, the RHI in AR seems to be experienced differently than in reality. As described in the previous chapter, this may have to do with the split in basis and augmentation, which is unique for AR as opposed to other media such as VR, but also a willingness to accept a virtual limb in AR, which does not actually exist, compared to plausibility to accept a rubber hand in reality. Therefore, in the following, I investigate both supernumerary hand and body illusions in AR, taking into account different variations in basis and augmentation. The Medical Research Ethics Committee of Utrecht had no objections to the execution of any of the following experiments.

4.5 Experiment 1: Required Multimodal Stimuli

The goal of this pilot study was to take a first step into understanding the fundamental multimodal criteria for the augmented reality supernumerary hand illusion (ARSHI). Here we give a concise description of the pilot experiment; an elaborate description can be found in (Gilbers, 2017). 27 participants wore a Meta DK1 HMD and experienced multimodal stimulation for three minutes, which was always a combination of different levels of visible number of hands, visuotactile synchrony, and active and passive visuomotor synchrony. After this, they filled out a questionnaire on arm ownership, agency, and hand-location; see Figure 4.1 for the setup of the experiment. Here we shall only discuss two methodological factors in particular: the appearance of the virtual hand, and the visuotactile stimulation by virtual smartwatch.

The virtual hand model was based on similar hand models that had been previously used in various RHI studies in VR (Slater et al., 2009). Although there have been indications that corporeality of the rubber/virtual hand could play a role in the illusion, for example when comparing a human-like hand with an abstract block (Tsakiris et al., 2010a), there was little support in the literature for the necessity of a more complex or higher quality virtual hand model in AR, therefore a corporeal hand shape with skin texture was deemed sufficient for this pilot, see the bottom row of Figure 4.1. Regarding the visuotactile stimulation, IJsselsteijn et al. (2006) mentioned various issues that affected ownership in both mediated conditions as opposed to the unmediated condition, one of which being sensory conflict after brushing a flat surface instead of an arm in the ‘MR’ case. To overcome such conflicts in this pilot experiment, visuotactile stimulation was implemented by adding a virtual smartwatch, which would flash with a notification screen, and synchronously/asynchronously vibrate on the wrist. Early iterations of this pilot study showed an improvement in compellingness of the visuotactile stimulation when using a smartwatch compared to a simple tapping mechanism.

Regarding the results, overall the reported ownership remained negative (below

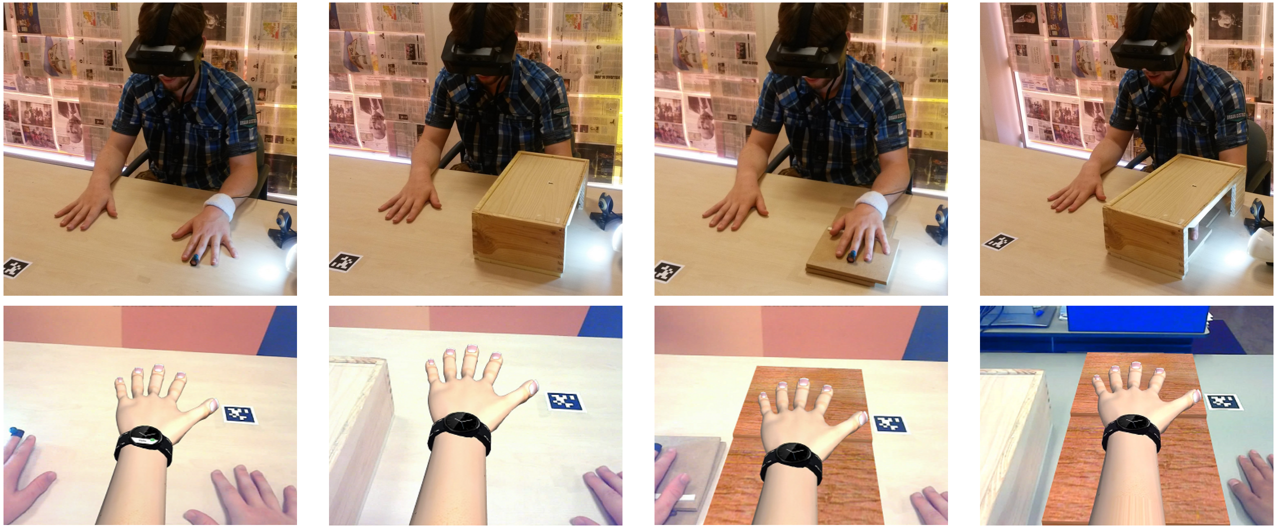


Figure 4.1: Setup of the pilot experiment for various conditions, from (Gilbers, 2017). A vibration motor is attached to the real left wrist, and the virtual left hand is wearing a smartwatch. A wooden box was used to hide the real hand in order to compare 'single' and supernumerary virtual arm conditions. (Top) front view, (bottom) participant's view; (left) visible real left hand with smartwatch notification and active/no movement, (left-center) real left hand not visible and active/no movement, (right-center) visible real left hand and passive movement, (right) real left hand not visible and passive movement.

midpoint) on a 7-point Likert scale, without clear trends of any factors. We suspect that the participants were possibly not inclined to accept the virtual hand as a plausible real hand due to its model-like appearance and the obvious visual dissonance between virtual and real. This nonacceptance of the hand could be a result of the a priori willingness factor mentioned earlier, and it is possible that a smaller gap between basis and augmentation may support acceptance. Lastly, agency occurred only when there was active synchronous visuomotor stimulation, and overall participants responded negatively to all hand-location questions, albeit less negatively to questions about a shift or a drifting feeling when only the virtual hand was visible than when both real and virtual hands were visible.

4.6 Experiment 2: Required Realism

A goal of this pilot study was to execute the first supernumerary body illusion in AR using an RHI paradigm, and to understand the role of body realism. Here we give a concise description of the pilot experiment; an elaborate description can be found in (Bommel, 2017). 34 participants wore an Oculus Rift with two mounted webcams two ensure stereoscopic vision. There were two factors: anthropomorphism (human body or block body) and visuomotor congruence. Participants watched an avatar standing two meters away from them, seen from the back, and moved the arms congruently (synchronous) or incongruently (prerecorded randomized movements), see Figure 4.2. After three minutes, the avatar was threatened and SCR was measured. After this, they filled out a questionnaire.

Both subjective and objective ownership results were greater in the congruent condition than in the incongruent condition, but overall the subjective results were again negative. The absence of an effect of anthropomorphism on body ownership was surprising. It was suspected that adding human movements to an avatar changes its overall realism: with the block body, the movements that clearly resembled arm movements made the avatar overall more human, whereas the human body combined with limited arm movements made the avatar less human. This resembles what was found by McDonnell et al. (2012), where face movements on a less realistic face was rated more pleasant than the same face movements on a more realistic face, resembling the uncanny valley effect. This would suggest that in order for any level of ownership to occur, only a realistic human body should be used, and if combined with other multimodal stimulation, this stimulation must at least match the level of realism implied by the visual appearance. Lastly, agency was rated positive for cases with active visuomo-

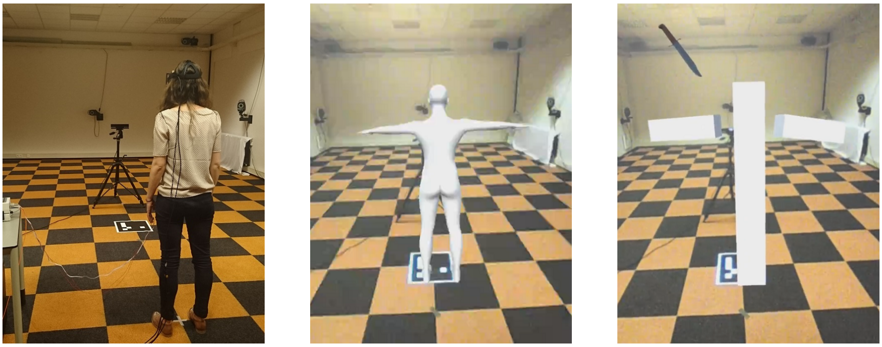


Figure 4.2: Views of the embodiment experiment. The virtual body appeared on a marker placed 2m in front of participant.

tor stimulation, and there were overall negative responses to all hand-location questions.

4.7 Experiment 3: Main Experiment

The goal of this study was to examine both the feasibility and requirements of the ARSHI. In the pilot experiments we found that hand realism in terms of appearance could be such a requirement, although it was not only due to corporeality as indicated by related work, but could additionally be associated with effects of other multimodal stimulation, and the distance between basis and augmentation. In the main experiment we examine this possibility by using a projection of the participant's real hand as the mediated third hand, and furthermore examine the induction of the illusion by means of realistic visuotactile and visuomotor stimulation.

4.7.1 Method

Six conditions were used in this experiment, which consisted of specific combinations of different levels of visible number of hands, visuotactile synchrony and visuomotor synchrony, see Table 4.1. It was expected that the number of hands would affect the level of experienced hand-location, not agency, and importantly also not ownership, showing that ownership over a supernumerary hand is as feasible as a 'single' hand illusion. Visuotactile synchrony would affect the level of ownership, but not agency or hand-location, and visuomotor synchrony would affect both the levels of ownership and agency, but not hand-location. These six specific combinations of factors were chosen rather than all possible combinations, in order to better inspect the feasibility of the SHI, as opposed to examining the relations between the various types of stimulation. Specifically, *C1* and *C2* compare the single virtual hand illusion with the SHI, while the other conditions compare necessary synchronous stimulation to asynchronous control stimulation. Remaining combinations, such as the those with both asynchronous visuotactile and asynchronous visuomotor stimulation, were excluded because they would not provide more insight on SHI feasibility than the included control conditions.

The supernumerary hand was chosen to be disconnected rather than connected in this experiment. Feuchtner and Müller (2017) indeed found that only the condition with connected virtual hand and invisible own hand resulted in any ownership, and that both disconnected and connected hand with own hand resulted in no ownership. Other studies without supernumerary hands suggest connectedness plays an important role in the experience of ownership (Tieri et al., 2015). However, in this experiment the goal was to let the supernumerary hand resemble a plausible own hand for the user by means of a 3D projection. It was possible that making this projection seem connected would result in an unwanted negative realism affect as was found in the pilot experiments.

The disconnected mediated left hand was added approximately in front of the participant, and in one condition a small wall was used to block the view of the real left hand, see Figure 4.3. Visuotactile stimulation consisted of a simplified virtual smartwatch face that would flash white for 250ms and vibrate for 250ms to resemble a notification, which would either be synchronous or asynchronous. There was a standard inter-notification onset interval of two seconds, and in the asynchronous case both the visual and tactile stimuli received a random offset between -500 ms and +500 ms, meaning in some occasions the visual part came first, and in others the tactile part came first. The visuomotor stimulation was always active movement in the form of a tapping motion of the left index finger, and also occurred synchronously or asynchronously, where the latter was implemented as a delay of one second. In the cases where there was no movement (*C3* and *C4*), the participant was instructed to keep their real hand still, and in the cases with movement (*C1*, *C2*, *C5*, *C6*) the participant was instructed to make their own lifting movements with their finger, and specifically to not copy the mediated hand.

For subjective measurement of ownership, agency, and hand-location, a questionnaire was used, see Table 4.2. This questionnaire consisted of nine questions that were answered on a 7-point Likert scale ranging from “strongly disagree”

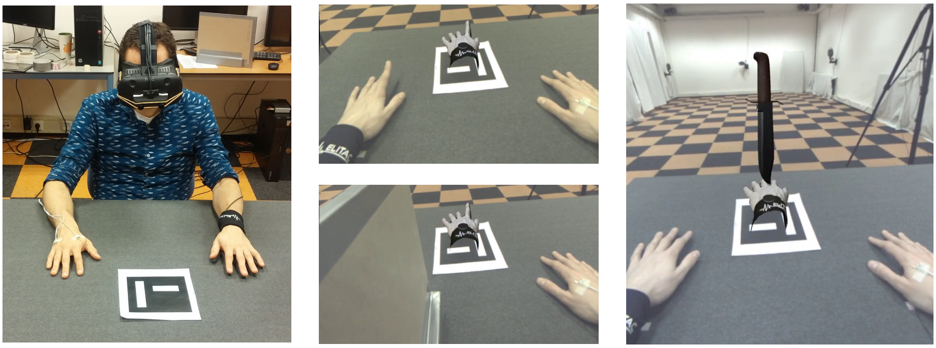
Table 4.1: The stimulation used in the six conditions of the main experiment. *C2* is a single virtual left hand condition where only the real right hand was simultaneously visible, whereas all others are variations of supernumerary hand conditions where all three real and virtual hands are visible. *C3* and *C4* did not include any hand movements, and *C5* and *C6* did not include vibrotactile smartwatch notifications.

Condition	*Hands*	*Visuotactile*	*Visuomotor*
C1	real and virtual	synchronous	synchronous
C2	virtual	synchronous	synchronous
C3	real and virtual	synchronous	-
C4	real and virtual	asynchronous	-
C5	real and virtual	-	synchronous
C6	real and virtual	-	asynchronous

to “strongly agree”. For objective measurement of ownership, SCR was recorded before and after a threat to the virtual hand, in the form of a virtual knife making stabbing motions towards the virtual hand, without it actually touching the surface of the virtual hand.

4.7.2 Material

A video see-through HMD was made by mounting an OvrVision Pro onto an HTC Vive, which has a 1080 by 1200 per eye resolution. The device was calibrated using the standard OvrVision Calibration tool. The application was developed in Unity 5.5.0 using SteamVR. The scripts were written in C# using Microsoft Visual Studio 2015. The participant’s left hand was recorded using a Kinect for Windows, and was displayed as if on top of a marker, see Figure 4.3b. The distance between the center of the real hand and the center of the marker was approximately 25cm horizontally (such that the virtual hand would appear in between the two real hands horizontally) and 25cm in depth. This way, both of the real hands and the virtual hand were in the field of view without requiring head movement. The participant wore an Elitac Tactile Display with a single vibrotactile motor fastened to the wrist by a black elastic band. The left hand was raised by approximately 8mm using a cardboard hand cut-out covered in the same cloth that covered the table, in order to bypass the limited depth resolution of the Kinect. Using a BioSemi ActiveTwo package, electrodes were attached to the right hand of the participant using a conductive paste. See Figure 4.3a for the front view of the participant. One computer, which could only be seen by the experimenter, ran ActiView to monitor the SCR, and another computer ran Unity. Because of problems that would occur after building, the experiment was run in play mode in Unity Editor, leading to a refresh rate of approximately



(a) Front view of experiment during certain condition.

(b) Participant's view during *C1* (top) and *C2* (bottom).

(c) Participant's view during the threat moment at end of *C1*.

Figure 4.3: Experiment setup. The participants were seated at a table with both arms extended in front of them. The left wrist was covered by an elastic band containing one vibration motor. The electrodes for the SCR reading were attached to the right hand. During stimulation participants would see an augmented mediated version of their left hand located above the marker.

Table 4.2: The list of questions that were asked after each condition, where the questions were presented in a random order, and answered on a scale from "strongly disagree" to "strongly agree". Depending on the condition, the formulation of *O3*, *L2* and *L3* was changed to concern either only 'vibrations' or 'movements'. Ownership is measured through three main questions that each reflect a different aspect of ownership: attribution, changing the body image, and the source of experienced sensations (Kilteni et al., 2012a; Guterstam et al., 2011), and the location questions cover three possible directions of change: multiple locations, full shift, and partial shift. Questions *O4* and *A2* were control questions. Although the supernumerary hand was mediated, the term "virtual" was used during the actual experiment for the clarity of the participant.

Q	*Question*	*Measure*
O1	It seemed as if the virtual hand was my hand.	Ownership
O2	It seemed as if I had three hands.	Ownership
O3	It seemed as if (1) the vibrations I felt were caused by the notification of the virtual smartwatch, and (2) the movements I felt were caused by the movements of the virtual fingers.	Ownership
O4	It seemed as if the real hand were slowly becoming digital.	Control
A1	It seemed as if I was controlling the virtual hand.	Agency
A2	It seemed as if the virtual hand was controlling my will.	Control
L1	It seemed as if my left hand was at two different locations.	Location
L2	It seemed as if I felt the vibrations and movements at the location of the virtual hand.	Location
L3	It seemed as if I felt the vibrations and movements somewhere between the real and virtual hand.	Location

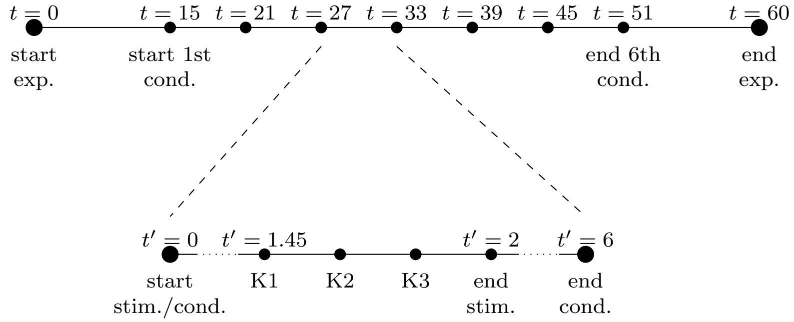


Figure 4.4: Outline of procedure for a single participant: (top) outline entire experiment, (bottom) outline single condition. t indicates the time within the entire experiment, t' the time within a single condition; both are listed in minutes. K1 indicates the moment the virtual knife appears floating above the virtual hand, K2 the moment it starts to make a stabbing motion, and K3 the moment it disappears (while the virtual hand remains visible).

35Hz. For general AR applications this rate is not ideal, and in our case was most noticeable during head movements. Since the participants were instructed to only focus on the three hands which were always in view, eliminating the need to move the head at all, we expect this hardly affected the experimental outcomes.

4.7.3 Participants

30 participants were recruited from the student pool of the university through advertisement by email and during lectures, with notification of a raffle for three 25 euro gift cards to an online department store as incentive. We remark that there was no student-teacher relation between the students and the experimenter/advertiser, therefore they were aware that not participating would not affect their course or education program progress. Average age was 22.7 (range 17-29, s.d. 2.8), 24 male, 6 female. 20 participants stated to know about or have experience with AR and 10 stated to know about or have experience with body illusions. It was confirmed that none of these factors influenced the results of this experiment.

4.7.4 Procedure

Participants were sent an information letter by email, including a copy of the consent form, at least a week prior to participation. This information letter included a verbal description of the threat they could expect during the experiment. Upon arrival in the laboratory, each participant was asked to read and sign the consent letter a final time, to wash and dry their hands, and take a seat at a table. Then the experimenter applied the electrodes to the distal phalanges of the index and middle fingers of the right hand. If for some reason this did not provide a signal or the participant had cuts or callouses on the fingers, the thenar and hypothenar eminences of the same hand were used. Next, the experimenter attached the vibrotactile motor to the left wrist, and measured the inter-pupilar distance (IPD), which in turn was adjusted in the HMD. The experimenter then put the HMD on the participant, and asked whether the participant could see the environment correctly; the IPD was then adjusted if necessary. Once the participant found this was the case, then the first session was started.

During each session, the participant was asked to experience a certain type of stimulation for 2 minutes, see Table 4.1. After 1 minute and 45 seconds of stimulation a knife would appear and float above the hand, while stimulation continued. After another 5 seconds, this knife would move in a stabbing motion towards the hand, and after another 5 seconds the knife would disappear and only the virtual hand remained. After a final 5 seconds, the virtual hand would also disappear and the session was over. After the session the participant would remove the HMD, and fill out a questionnaire on a tablet where the questions were presented in random order, see Table 4.2. After approximately 4 minutes, the next session was started. Each session including break would therefore take approximately 6 minutes.

After all six conditions were completed, which occurred in a unique random order for each participant, the participant was asked a few post-experiment questions to check for lasting effects of the experiment. These were based on those used by Kilteni et al. (2012b). Lastly, the participant was explained what the purpose of the different conditions was, and was allowed to ask any questions they had. They were also asked whether they would like to take part in the raffle. An outline of the portion of the experiment that took part in the laboratory can be found in Figure 4.4.

At least three days after participation, each participant was called and asked a second round of post-experiment questions. In two cases where the participants could not be reached, the questions were asked by email. There were no lasting effects for any participants.

4.7.5 Results

Questionnaire Data

For each question a Friedman test was used to test for an overall effect over conditions in SPSS. All questions except for *A2* resulted in a significant effect, and post hoc analyses with Bonferroni correction were run. See Figure 4.5 and Table 4.3 for the results. Generally, the responses for the conditions with both visuotactile and visuomotor synchrony were highest, namely *C1* (real and virtual left hands, synchronous VT, synchronous VM) and *C2* (only virtual left hand, synchronous VT, synchronous VM). The conditions with any asynchronous stimulation showed the lowest responses, namely *C4* (real and virtual left hands, asynchronous VT, no VM) and *C6* (real and virtual left hands, no VT, asynchronous VM).

Participant Comments

Of the 180 possible post-condition comments and the 60 possible post-experiment comments, there were a total of 69 and 48 non-empty meaningful comments, respectively. Each of these comments was labelled, where each comment could receive more than one of 19 labels, and the labels ranged from 'immersive experience' ("...the hand felt completely real") to 'meta experiment' ("What about a real knife instead of a virtual knife?"). Notable cases were that 14 participants remarked that the hand did not look realistic enough in a post-condition comment, 27 participants remarked positively regarding the experiment in general in a post-experiment comment, 10 participants remarked experiencing some form of discomfort in a post-experiment comment, and 8 participants remarked an experienced imbalance over modalities in a post-condition comment.

SCR Data

The SCR values were computed using a difference function: take the average of the 10 seconds before the threat appeared and subtract this from the maximum level that occurs in the 10 seconds that the knife is present (5 seconds static knife presence, and 5 seconds stabbing motion), and finally add 1 and take the logarithm (base 10); see Figure 4.6 for an example of skin conductance during these 20 seconds. A Friedman test over all the SCR data showed there was no significant effect of condition, $\chi^2(5) = 2.636; p = 0.756$. However, using a post-hoc Friedman test on the same SCR values, but now arranged by order of performed condition, there was a significant habituation effect towards the

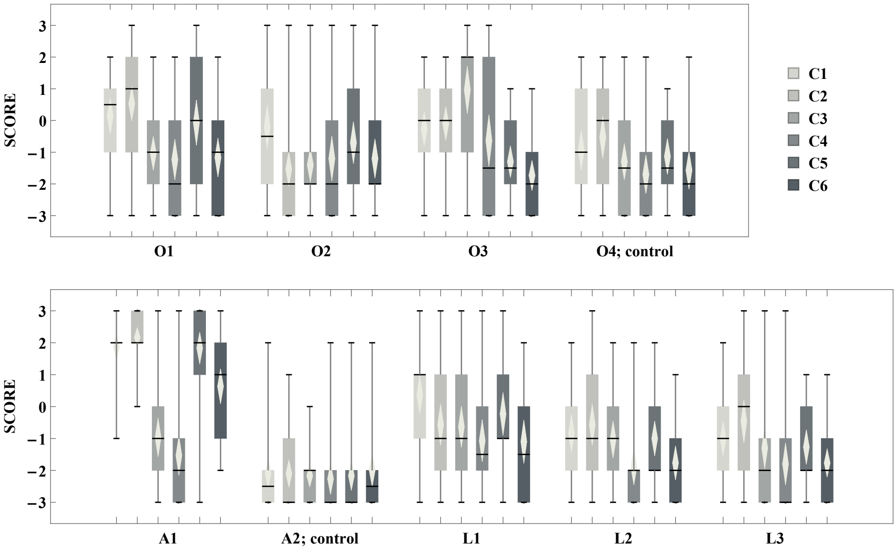


Figure 4.5: Boxplots of all questionnaire responses, on a seven-point Likert scale where -3 corresponds to "strongly disagre" and +3 to "strongly agree". The medians are marked with a black horizontal line, and the means +/- mean standard errors are marked as light grey diamonds.

Table 4.3: The results of Friedman tests per questions, and the accompanying post hoc tests with Bonferroni correction. All Friedman test results are asymptotic, and all Wilcoxon signed-rank test results are exact.

Q	$\chi^2(5)$	p	post-hocs at $p < 0.00333$
O1	62.827	<0.0001	*C1*,*C2*,*C5* > *C3*,*C4*,*C6* with exception *C3* = *C5*
O2	19.435	0.002	*C1* > *C2*,*C3*
O3	48.653	<0.0001	*C1*,*C2* > C5,*C6* and *C3* > *C1*,*C4*,*C5*,*C6*
O4	25.265	0.0001	*C1*,*C2* > *C4*
A1	102.327	<0.0001	*C1*,*C2* > *C3*,*C4*,*C6* and *C5*,*C6* > *C3*,*C4* and *C5* > *C6*
A2	3.383	0.641	-
L1	23.582	0.0003	*C1* > *C4*,*C6*
L2	34.694	<0.0001	*C1*,*C2*,*C3* > *C4*
L3	24.482	0.0002	*C2* > *C6*

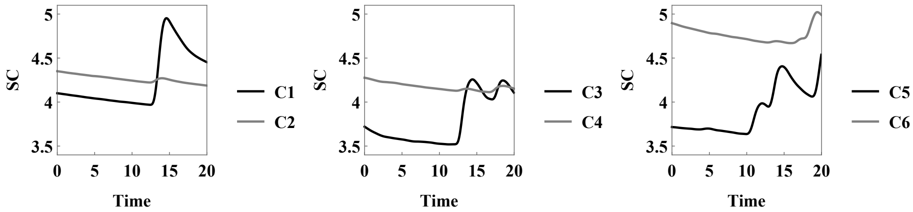
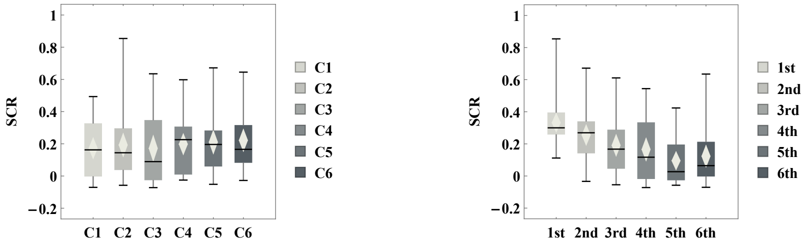


Figure 4.6: Representative skin conductance (SC) plots of one participant for each condition. The vertical axis represents SC in microSiemens, the horizontal axis time in seconds. The single SCR value used for analysis was computed by taking the average of the first 10 seconds of the SC curve (baseline) and subtracting this from the maximum level that occurs in the last 10 seconds of the SC curve (knife presence), with finally a log transformation log(value +1).

threat, $\chi^2(5) = 38.621; p < 0.0000003$, see Figure 4.7. Using multiple Wilcoxon Signed Ranks Tests with Bonferroni correction, the differences were that the *first* presented condition was significantly greater than the *third*, *fourth*, *fifth* and *sixth* (all $p < 0.0003$), and *second* was significantly greater than *fifth* and *sixth* (all $p < 0.001$).

Relation between Questionnaire and SCR Data

Since the questionnaire data regarding ownership *did* result in significant effect while the SCR data did *not*, it is important to evaluate how participants responded individually post-hoc. To do this, we compare each participant's SCR 'character' to their questionnaire response 'character'. Here, the SCR character is defined as the range in which a participant's SCR responses lies; similarly, the questionnaire response character is defined as the range in which the responses to *O1*, *O2*, and *O3* lie, which are the three positively formulated non-control questions regarding ownership. We can approach these ranges by performing a cluster analysis on the means and standard deviations of both questionnaire and SCR data, resulting in an analysis of a 4-dimensional space. Considering the number of data points, hierarchical clustering was chosen for the analysis on the four variables: SCR_means, SCR_st.devs., Q_means, and Q_st.devs. Clustering with between-groups average linkage, squared Euclidean distance, Z-score normalization, and by applying the stopping rule in the agglomeration schedule, we find the clustering as shown in Figure 4.8. We see that besides the four possible outlying points (cluster numbers 3, 4 and 5), there are two large clusters: the participants with low questionnaire response means and low standard devi-



(a) SCR by condition. **(b)** SCR by performed order.

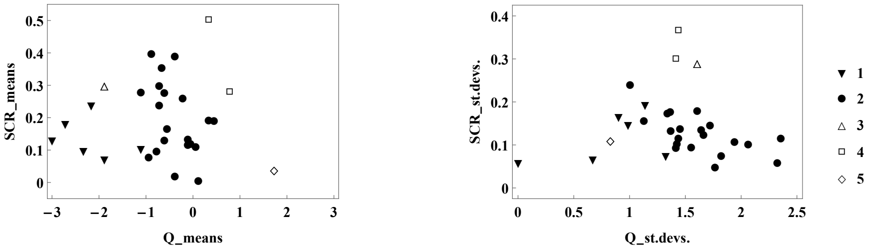
Figure 4.7: Boxplots of the SCR values. Each value is computed by subtracting the average of the 10 seconds before the appearance of the knife from the maximum value in the 10 second of knife presence, then adding 1 and taking the log with base 10, i.e. log (value + 1). The medians are marked with a black horizontal line, and the mean +/- mean standard error are marked as light grey diamonds.

ations who also have low to medium SCR means (cluster number 1), and those with high questionnaire response means and high standard deviations who also have low to high SCR means (cluster number 2). Executing the analysis for the remaining three variables, i.e. all but SCR_st.devs., provides the same clustering. To show these variables were not interdependent, a multiple correlation test with SCR_means as dependent variable and Q_means and Q_st.devs. as independent variables was executed, and did not result in a significant correlation, $F(2, 29) = 0.160$, $p = 0.853$ with $R^2 = 0.012$.

4.7.6 Discussion Ownership Results

Here, I shall briefly discuss the subjective ownership results with respect to the pilot experiments; the main discussion can be found in Section 4.9. The questionnaire results of the main experiment demonstrate that the SHI is indeed feasible in AR. Generally, the conditions with synchronous feedback led to positive experiences of ownership, regardless of number of visible hands, and those with asynchronous feedback led to negative experiences of ownership.

Although the mediated hand projection *was* adequate in the main experiment compared to the pilots in terms of resulting in an overall positive ownership experience, there were still many comments on the bad hand appearance, such as that the virtual hand “looked a bit deformed”, “was not super realistic”, and “looked weird.” One participant explained that although the virtual hand was obviously a projection of their own real hand, it was not *really* their hand. These



(a) Clustering on means.

(b) Clustering on s.d.s.

Figure 4.8: Cluster graphs after clustering on four variables: questionnaire response mean and standard deviation over *O1*, *O2*, and *O3* over all conditions, and SCR mean and standard deviation over all conditions. Two graphs are used to display the five resulting clusters on all four variables, and we emphasize that both graphs represent the same clustering result.

contrasts show that another factor may be at play in the sense of embodiment. As introduced in Chapter 3, in AR it could be a matter of passing a barrier of belief that is emphasized by the dissonance between basis and augmentation, which is in turn possibly determined by a priori willingness to accept augmented limbs. Willingness is often mentioned in the context of general media use. For example, by drawing upon phenomenological similarities between VR and narrative text, Ryan (1999) discusses the necessity of Coleridge's willing suspension of disbelief in order to regard non-actual possible worlds as actual. Translating Coleridge's views to VR and AR, Ryan argues that if the creator of the virtual/mixed environment ensures that the environment is as truthful as possible, then the user would willingly accept the fake environment as real. That is, the user actively decides to accept the fiction if it seems plausible, which differs greatly from, for example, the ideas on how presence passively occurs: without choice as a result of, for example, system variables (Lombard and Ditton, 1997). Still, willing suspension of disbelief may be related to immersive tendency, and in turn presence, in the sense that an individual with high immersive tendency may have those personality traits that make them more prone to decide to accept fiction, as suggested by Weibel et al. (2010).

Willingness is an uncommon concept in RHI literature, although Stone et al. (2018) informally describe a form of suspension of disbelief: "a fundamental difference between 'feeling' and 'knowing' during the rubber hand illusion that must be overcome to experience the illusion. For example, in the RHI, the individual must override the knowledge that the rubber hand is not his in order to surrender to the feeling that it is, encouraging the incorporation of a foreign hand into the

Table 4.4: Shortened version of the Immersive Tendency Questionnaire, ITQ-short, by Weibel et al. (2010).

Q	Question
IT1	Do you ever become so involved in a movie that you are not aware of things happening around you?
IT2	Do you ever become so involved in a TV program or book that people have problems getting your attention?
IT3	Do you ever become so involved in a daydream that you are not aware of things happening around you?
IT4	Do you ever have dreams that are so real that you feel disoriented?
IT5	Have you ever gotten scared by something happening on a TV show or in a movie?
IT6	Have you ever remained apprehensive or fearful long after watching a scary movie?
IT7	How good are you at blocking out external distractors when you are involved in something?
IT8	Have you ever gotten excited during a chase or fight scene on TV or in the movies?
IT9	Do you ever become so involved in doing something that you lose all track of time?

sense of the bodily self". In the next experiment, I investigate whether immersive tendency may be related to subjective ownership in the AR supernumerary hand illusion.

4.8 Experiment 4: Immersive Tendency

4.8.1 Participants

All participants from the main experiment were contacted to reparticipate in this follow-up experiment, 17 months after the main experiment. Of the 30 participants, 2 no longer had the same contact information and could not be reached; 23 of the remaining participants consented to take part; age mean 22.3, s.d. 2.3 (at the time of the main study); 5 female, 18 male. It was confirmed that age and sex did not effect the results of this study.

4.8.2 Material

A shortened version of the original Immersive Tendency Questionnaire (ITQ) by Witmer and Singer (1998) was used, which is based on the analysis performed by Weibel et al. (2010) (from here on ITQ-short), see Table 4.4. These questions were answered on a 5-point Likert scale. We used the short version of the immersive tendency questionnaire, because the original long version had been constructed on a theoretical basis, lacking statistical validation.

Regarding the ownership data, we used part of the subjective data from the main experiment, namely the responses to the most direct ownership question “It seemed as if the virtual hand was my hand.” (*O1*) for the conditions *C1* (real and virtual left hands, synchronous visuotactile, synchronous visuomotor) and *C2* (only virtual left hand, synchronous visuotactile, synchronous visuomotor). These conditions were chosen because they are theorized to create a sense of ownership, and the results showed high variance among participants; that is, they lead to positive and negative ownership responses. These two conditions were not statistically significantly different for the original 30 participants in the previous study.

Although also theorized to create a sense of ownership, we did not include the results of conditions *C3* (two real hands and virtual hand, visuotactile synchrony) or *C5* (two real hands and virtual hand, visuomotor synchrony) in our analysis because those ownership results were overall weaker. We also did not include SCR data because the main study did not find the expected differences between conditions, and the results may have been driven by a habituation effect. Also, the main study concerned the sense of embodiment (ownership, agency, and hand-location) and thus included multiple questions on all three concepts. This follow-up study only investigates explicit ownership, not related ownership experiences of changing the body image (*O2*) or the source of experienced sensations (*O3*), thus the results of these questions are not included here. Similarly, the responses to the questions on agency and hand-location are also not investigated.

4.8.3 Procedure

All previous participants were sent an information and consent letter approximately 17 months after the execution of the main experiment. If they consented, they were digitally sent the ITQ-short one week later, that they could fill in in their own time outside of the laboratory. This approach was chosen to ensure a high rate of reparticipation. Each participant could take part in a raffle for one 10 euro gift card to an online department store. The questionnaire was closed nine days after sending, and no more responses were recorded. All participants who consented responded within this time. The raffle was awarded five days after the questionnaire was closed.

4.8.4 Results

All responses to the ownership questions and immersive tendency questions were ordinal. For this reason, an ordinal regression was performed in R with factor C (condition; 2 levels) and ordered factor IT (immersive tendency median; 3 levels since only values 2, 3, and 4 occurred) without interaction variable.

The assumption of proportional odds was upheld for C ($p = 0.8104$) and IT ($p = 0.0718$). The model was significant ($\chi^2(3) = 14.298, p = 0.0025$), Nagelkerke pseudo $R^2 = 0.276$. C showed no statistically significant differences ($\chi^2(1) = 1.4673, p = 0.2258$), and IT did ($\chi^2(2) = 12.8166, p = 0.0016$) with parameter estimate $\beta = 3.3648$ ($z = 3.119, p = 0.00182$). The estimate indicates that if a participant shows one unit increase in IT (e.g. from 2 to 3), then we expect a 3.3648 increase in the ordered log odds of experiencing a higher level of ownership, assuming all other variables in the model are held constant. The post-hoc tests with Tukey adjustment showed that IT=2 (ownership median= -0.5) and IT=3 (median= 0) differed significantly from IT=4 (median= 2) ($p < 0.05$). See Figure 4.9 for an overview of the data.

4.9 Discussion

The questionnaire results of the main experiment demonstrate that the ARSHI is indeed feasible. Regarding the ownership questions, the conditions with multiple synchronous stimulation (i.e. both visuotactile and visuomotor) resulted in a positive experience of attribution (*O1*), and asynchronous conditions led to a negative feeling of attribution. Unexpectedly, both single synchronous conditions (i.e. either visuotactile only or visuomotor only) were experienced as weak. Curiously, only the multiple synchronous stimulation case with three visible hands led to a moderate experience of having three hands (*O2*), while all other conditions, even those with three visible hands, did not cause this experience. This feeling was most negative for the case with only two visible hands, indicating a higher weighting to visuals in the context of changing the overall body image throughout the experiment. Regarding the source of the sensations (*O3*), synchronous visuotactile stimulation alone showed the highest response, which is

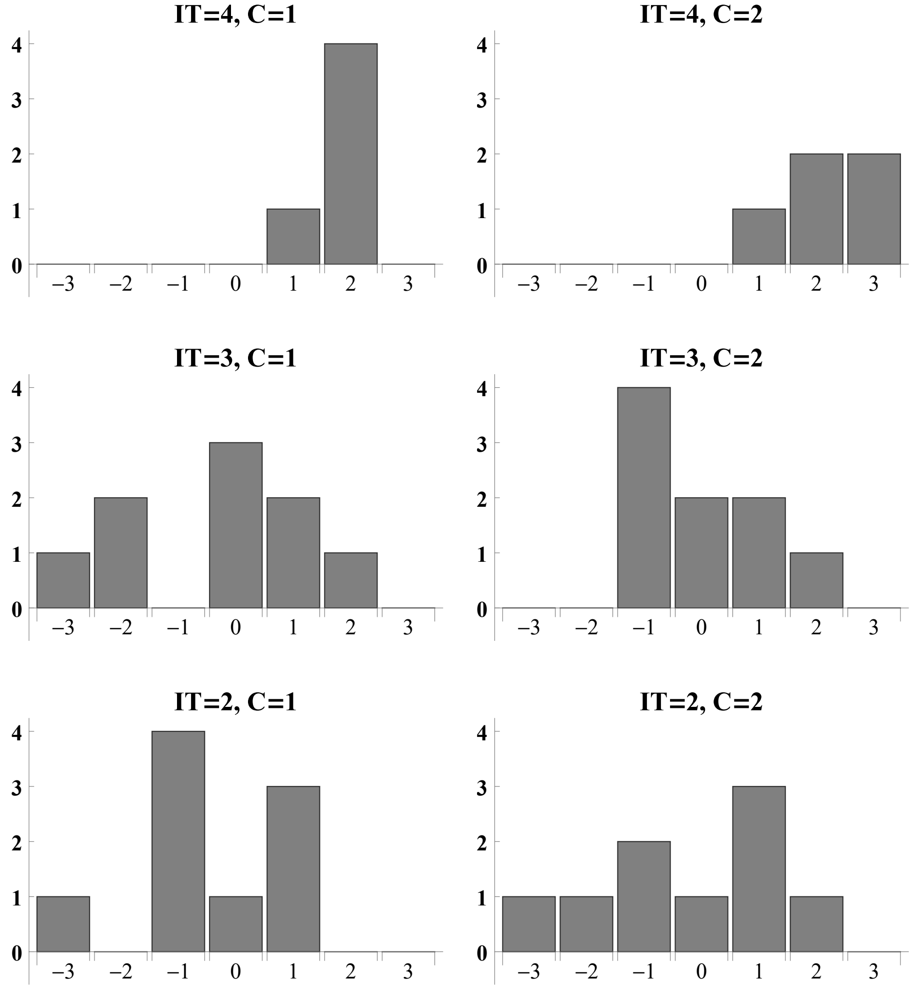


Figure 4.9: Histograms of the ownership responses, categorized by immersive tendency (IT) and condition (C). For simplicity, the 7-point Likert scale ownership data "strongly disagree", . . . , "strongly agree" is converted to -3 , . . . , 3.

surprising considering the weak attribution results for this condition. The participant's comments on imbalance over modalities often concerned the multiple synchronous stimulation conditions, where it felt as if the visuotactile stimulation had more of an impact than the visuomotor stimulation, but this is not reflected in the results of the ownership questions. Had this been the case, then the multiple synchronous stimulation conditions would have equal responses as the single synchronous visuotactile condition. All conditions showed negative experiences of the real hand becoming digital (*O4*; control). We remark that although this question was intended as a derivation of the original "turning rubbery" question by Botvinick and Cohen (1998), the "turning digital" version could have been interpreted differently. While both terms could reflect a materialistic aspect, 'digital' could also reflect a visual-only aspect, namely whether the arm seemed to be made up of pixels. Since a video see-through HMD was used here, participants could have been inclined to respond less negative than the rubbery counterpart in a classic RHI; however, we do not expect that this occurred since the results do not differ strongly from those of the synchronous visuotactile condition reported by Botvinick and Cohen.

Both multiple stimulation conditions resulted in the highest sense of agency (*A1*), and the single synchronous visuomotor stimulation followed closely. Approximately half of the participants also reported a positive sense of agency in the asynchronous visuomotor condition, which is consistent considering they still had control, albeit delayed. As expected, the single visuotactile conditions resulted in negative feelings of agency, as did the control question of whether the mediated hand was in control (*A2*; control) in all conditions. For the experience of arm-location, most participants reported feeling multiple left hand locations in the multiple synchronous stimulation condition where three hands were visible (*L1*), and this sensation was negative for most participants in all other cases. Regarding the question of whether the sensations were felt at the location of the virtual hand (*L2*), all results were negative; however, the conditions with at least synchronous visuotactile stimulation were least negative. Finally, regarding whether the sensations were felt between the real and virtual hand (*L3*), we see a weak response for the condition with only two visible arms; all other conditions were negative. Together these hand-location results indicate that when there is multiple synchronous stimulation, if two left hands are in view then two left hand positions are experienced, as opposed to an experience of drift when there is only one left hand in view.

These questionnaire results concur with related work considerably. With respect to SHIs for example, we achieved similar response levels to a variety of questions used in Guterstam et al.'s experiments 1 (compare to *C3* and *C4*) and 5 (compare to *C1* and *C2*, with the exception of synchronous visuomotor stimulation). For the RHI in AR, again similar response levels are found in Suzuki et al.'s (2013) tactile condition (compare to *C2*, with the exception of synchronous visuomotor stimulation) and IJsselsteijn et al.'s (2006) VR condition (compare to *C2*, with the exception of synchronous visuomotor stimulation). Notably, higher ownership results were found in our study compared to the disconnected virtual hand and visible real hand conditions in Feuchtner and Müller's (2017) study (compare to *C1* and *C2*, with the exception of synchronous visuotactile stimulation).

Although further comparison with more studies from Sections 4.1, 4.2 and 4.3 is hardly possible due to large differences in experimental setup, the conclusions are similar: it is possible to experience a supernumerary limb as one's own alongside the real hand in AR. Nevertheless, a few remarks on methodological differences between these studies and this study must be made. Firstly, it should be noted that these studies focused only on the concept of ownership, and their questions therefore intend to measure ownership only. *L2* and *L3* are examples of questions that are regarded as ownership questions in other experiments, but in this experiment are considered hand-location questions because they explicitly ask about the location of certain experienced sensations. The study by Guterstam et al. regards *L3* as an ownership control question, indicating this may indeed not measure ownership in the first place. Secondly, agency was not measured in any of these studies, and is a contribution of this study to the existing body of literature. Specifically, according to the questionnaire results of the main experiment, active synchronous visuomotor stimulation led to a sense of agency over the virtual hand, while active asynchronous stimulation did not, regardless of the visual presence of the real hand, as expected. Thirdly, although we have discussed visuotactile stimulation as a general approach, the first pilot and main study used a virtual smartwatch rather than a more common (virtual) brush stroking. Since the responses in this and other studies are similar, we do not expect that this difference affected the experience. There is also no reason to believe that brushing should be considered 'more correct' than a smartwatch; Botvinick and Cohen did not explain or defend their initial choice for brushing Botvinick and Cohen (1998). By contrast, the limited movement freedom may have influenced the ownership results: the subjective ownership results for the movement condition in Suzuki et al. (2013) are slightly higher than the results of our movement condition without tactile feedback, indicating that in fact more ownership could have been experienced with more movement freedom. Although this does not influence the major conclusions of our study, namely that the SHI is feasible in AR, it does provide a fruitful area for future research.

4.9.1 Questionnaire versus SCR

It is puzzling why the questionnaire results in the main experiment show such clear differences, while these same differences are absent in the SCR data. The SCR results also do not seem to reproduce similar SCR results from other studies.We suspect that the habituation effect that occurred in this experiment was stronger than any possible effect based on condition. It is noteworthy that, although SCR habituation is a very basic, common, and strong effect, it is, to our knowledge, not mentioned in any RHI studies that use it as an objective measure. We suggest that future studies critically evaluate habituation effects. We remark that there is so far no evidence for a similar habituation effect in the questionnaire data. By defining questionnaire habituation as less variation in the responses over time, it can be examined by comparing the standard deviations of all questionnaire responses split by order. The absence of a habituation effect in the questionnaire responses was subsequently confirmed, $F(5, 145) = 1.699; p = 0.139$.

To relate the questionnaire and SCR data, a cluster analysis was executed. There were two main groups among the participants: a group with overall low but consistent questionnaire responses with low to medium SCRs, and a group with overall high and more inconsistent questionnaire responses with low to high SCRs. I elaborate why the choice of cluster analysis over, for example, linear regression, is justified. Prior to this experiment, one could have reasonably hypothesized that participants with low SCR responses would also respond more negatively to the questionnaires (and analogous for high and positively). The results of the main experiment, however, showed that participants did not react in this way: the significant differences that appeared in the questionnaire data did not also appear in the SCR data. To find out whether certain links could be deduced from the different data types, a linear regression would have been insufficient, since it is clear that the questionnaire and SCR data would not lie on the hypothesized line. Instead, we were inclined to search for possible groups of participants that defined regions in the 4-dimensional search space. The multiple correlation test was furthermore not significant; however, this is not surprising since the questionnaire mean and standard deviation seem to predict the range within which the SCR mean occurs, not necessarily the SCR value itself.

4.9.2 Immersive Tendency and a Model for Ownership in AR

Many works have been dedicated to explaining which processes lead to ownership experiences; see Tsakiris (2010) for a review of existing models. These models generally structure various neurocognitive bottom-up and top-down processes, but do not extensively describe how variables that are related to large variances fit into these structurings. These existing models also do not differentiate between mediated (i.e. in VR and AR) and non-mediated versions of the illusion, while a model for the mediated version may have components related to media use. In this discussion, I propose a two-level processing model for mediated instances of the RHI that accounts for commonly occurring variation across participants, by considering the illusion not only a neurocognitive phenomenon in the first level, but also a media experience in itself in the second level. I propose that a large portion of the variance arises outside of the bottom-up and top-down processing mechanisms, but occur afterwards as a result of perceptual hypothesis testing, where certain user variables related to the media use come into play.

The motivation for making hypothesis testing explicit in this proposed model originates from a model of spatial presence by Wirth et al. (2007). See Figure 4.10 for an illustration. Starting with bottom-up and top-down information as input, the first level processing outputs a tentative limb model. The first level processing can be influenced by certain user and media variables. For example, (proneness to) certain mental disorders (Asai et al., 2011; Germine et al., 2013; Kállai et al., 2015), could be a *user variable for construction.* Regarding *media variables for construction*, one could imagine that technological properties such as image quality or system fidelity could play a role here. In the second level, the tentative limb model is taken as a hypothesis for perceptual hypothesis testing. This processing could again be influenced by certain user and media variables, and the outcome, after repeated acceptance of the tentative model, is then ownership. The results of Experiment 4 suggest that immersive tendency could be such a user variable for testing.

Here we explain the differences and similarities between this proposed model and existing ownership models. In the model proposed by Tsakiris and Haggard (2005), bottom-up stimuli are only registered in the parietal area 5 neurons when there is enough multimodal congruence between the fake and the real hand. If there is then also affirmation from top-down influences, this population of neurons is activated, and can result in the sense of ownership. In these cases, there is no differentiation between the positive registration of the stimuli, while considering existing self-representations, and the development of the sense of ownership. Putting this in terms of our proposed model, this would mean that a new hand model is formed as ownership occurs. According to Maselli and Slater (2013), ownership would not occur when the top-down influences are incorrect, even when the bottom-up stimuli are correct. Again putting this in terms of our proposed model, this becomes an interesting case: is a hand model formed, but

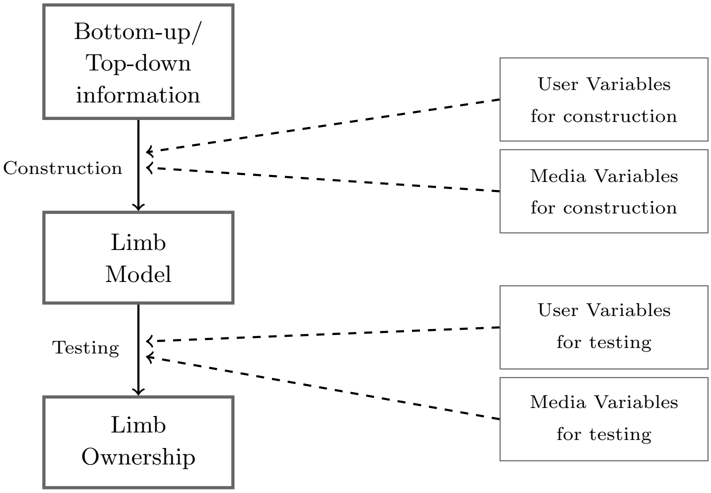


Figure 4.10: Illustration of the proposed two-level model for ownership. The first level concerns processing of bottom-up and top-down influences and results in the construction of a tentative limb model. The second level concerns testing of the hypothesis whether the limb is owned. Both levels are influenced by a variety of user and media variables.

without ownership, or is there no model and also no ownership? We argue that the former is the case: even in experimental conditions where there is no ownership but bottom-up stimuli are congruent, participants are always aware of a suggested hand model, and simply do not accept it due to other inconsistencies, and possibly other user variables. It should be pointed out that we are not showing that these existing models are incorrect. Indeed, we agree that in our proposed model the bottom-up and top-down processing occurs at least similarly to what is described in existing models. Instead, we make one particular processing level explicit, that occurs implicitly in existing models. Therefore, previous results that are explained by these existing models should also be explained by our proposed model.

4.10 Conclusion

In this chapter, I further examined the interesting AR avatar embodiment scenarios listed in Chapter 3. Through a series of experiments, I showed that one can embody a supernumerary hand alongside the own body in AR, but only when the realism of the hand appearance matches the realism of the multimodal stimulation, and when the gap between basis and augmentation is small. It is possible that similar results may be found in relation to supernumerary bodies with these requirements. In the next chapter, I examine the practical benefits of the sense of embodiment.

Chapter 5

Supernumerary Embodiment and Task Performance

With the increasing use of virtual technologies in RHIs, the question arises whether experiencing a sense of embodiment should be a goal of the simulation in itself, or whether it can be used as a means to achieve other goals, such as eliciting certain emotional responses (Waltemate et al., 2018) or altering motor behaviour of virtual body parts (Burin et al., 2019). For example, practical benefits would include a relation between embodiment and motor behavior performance. Such a relation would expand and further strengthen the application of VR and AR in domains beyond entertainment such as healthcare, education, and teleoperations. Recently, it has been postulated that embodiment of a remote manipulator can improve dexterous performance, based on evidence from VR and prosthesis use (Toet et al., 2020). The reasoning is that an enormous amount of research has been focused on improving the transparancy of the teleoperation system, since any technical flaw may hamper the execution by an operator. From this one could argue that if a system seems transparent to the user, it implies that the robot body is experienced as the operator's own body. A number of recent studies examine this postulate through only body ownership and have found mixed results. For example, Grechuta et al. (2017, 2019) found clear positive correlations between body ownership and performance in various tasks in VR. In contrast, Shin et al. (2021) found that greater body ownership may cause an increased risk perception, which in turn leads to degraded performance in pick-and-place tasks in VR. Notably, these works only exploit the use of VR technology, but do not study the same effects in AR, so it remains unclear whether extending the body with a virtual hand in AR rather than replacing the real hand by a virtual hand as is typical in VR, may affect any possible relation between ownership and performance.

Interacting with objects through virtual hands in AR requires continuous action, in which case there is evidence for a relation between the sense of agency and task performance (Wen et al., 2015). The link, then, between the sense of ownership and performance may be through the sense of agency. If taking into account the evidence that a sense of ownership may facilitate a sense of agency (Kalckert and Ehrsson, 2012), one may suggest that agency is a possible mediator of the relation between a sense of ownership and performance. A small number of studies in VR have measured these three phenomena simultaneously (Egeberg et al., 2016; Laha et al., 2016), but it is not common that the role of agency is further studied. In this chapter, I investigate the relation between the sense of ownership and task performance in an AR task. Specifically, the objective is to investigate whether this relation is established through the sense of agency in an ARSHI.

5.1 Ownership and Agency

Research on the possible link between ownership and agency has mixed results, see (Braun et al., 2018) for a thorough review. Here, I highlight a few recent studies. Tsakiris et al. (2010b) used fMRI to compare activated brain areas in an RHI paradigm, with factors movement (active or passive) and visual feedback (synchronous or asynchronous), measured through questionnaires. Their questionnaire responses to ownership and agency over the rubber hand both followed patterns typical to other RHIs, supporting an additive model of agency to ownership (i.e. agency entails body ownership). The neuroimaging data, on the other hand, showed that ownership and agency were associated with distinct and exclusive patterns of activation, supporting an independence model (i.e. they are qualitatively different experiences). The authors argue that the inconsistent results could be explained by the participants using common sense while responding to the questions, and that there may also not be a one-to-one mapping between brain activity and conscious experience.

Another study supporting the independence model is that of Kalckert and Ehrsson (2012). Here, the authors performed a series of four experiments in an RHI paradigm to simultaneously measure ownership through proprioceptive drift and questionnaire and agency of the fake hand by questionnaire. In the first pair of experiments they used factor movement timing (synchronous or asynchronous) while the rubber hand was passively moved, and found that both ownership and agency were experienced in the synchronous condition, and in the asynchronous condition there was lower but positive agency and no ownership. They found no correlation between proprioceptive drift and agency in either condition. In a second pair of experiments, they used factors movement mode (active or passive) and hand position (congruent versus incongruent) with synchronous movement. The questionnaire responses found a double dissociation of the two experiences: there was strong ownership (in both measures) and agency in the active congruent condition, ownership but no agency in the passive congruent condition, agency but no ownership in the active incongruent condition, and neither in the passive incongruent condition. Ownership and agency statements were positively correlated in the active congruent condition, but not in the other conditions. In summary, the results suggested ownership and agency were independent processes, and ownership modulated agency, that is, stronger agency was experienced when the hand model was owned.

A study supporting the additive model is that of Burin et al. (2017). The authors combined an RHI paradigm with a sensory attenuation (SA) paradigm to examine how body ownership contributes to agency, by only using self administered shocks (and not also by an ‘other’ as is typical in SA-only experiments) in three ownership-related conditions. In the synchronous visuotactile condition where ownership ratings and proprioceptive drift were high, intensity ratings were low and agency questionnaire responses were high, meaning the movement of the fake embodied finger was subjectively misattributed to the participant’s own will and the stimulus intensity delivered by that finger was attenuated. On the other hand, in the two conditions where ownership ratings and drift were low (asynchronous visuotactile, and synchronous visuotactile with incongruent hand position), intensity ratings were high and agency questionnaire responses were low. This means that the movement of the fake not-embodied finger was not misattributed to the participant’s own will and the stimulus intensity delivered by that finger was not attenuated. In summary, in the absence of (intent of) motor actions, when participants experienced ownership they also experienced agency, and when there was no ownership there was also no agency. The authors conclude: “owning the body would lead to the inference ‘since this is my body part, any action would be intended by me’ ”.

Lastly, in a similar study to that of Burin et al., Pyasik et al. (2018) perform two separate experiments on the same group of participants, one measuring ownership over a fake hand in an RHI through proprioceptive drift and questionnaire, the other measuring experienced agency in Libet's clock paradigm through intentional binding, intensity attenuation and questionnaire. Both experiments had the typical result patterns. These results were subsequently examined for correlations, and the authors found a positive correlation between proprioceptive drift and attenuation, and importantly, no correlation between both questionnaires. These results are in contrast to the study by Tsakiris et al. discussed above, in that here the questionnaire responses showed no overlap, whereas the quantitative measures did. The authors explain this may be because that study and many more used movement of an embodied fake hand to examine both ownership and agency. They conclude that spatiotemporal constraints in integrating sensory-related signals are common to both body ownership and sense of agency, supporting an additive model, whereas their subjective experience would rely on additional processes specific for any given sense, supporting an independence model.

To summarize, the discrepancy of results concerning the link between agency and ownership seems to depend on the type of measure, where qualitative measures typically find an overlap in experiences and quantitative measures do not, and whether movement was used to elicit agency, which may also accommodate body ownership. For VR and AR, the use of questionnaires combined with movement to examine body ownership and agency are in the majority compared to other measures and setups, meaning there may be a bias towards an additive model in the literature. Nonetheless, since the purpose of this study is to examine performance in a sensorimotor task which is to be executed by moving limbs, I hypothesize that there will be a positive correlation between experienced ownership and agency of the virtual hand (*H1a*).

RHI related studies in AR are rare, thus it is not straightforward whether factors from reality and VR may also influence experienced ownership and agency in AR in a similar manner. In the main experiment from Chapter 4, varying ownership experiences seemed to rely mostly on increasing numbers of synchronous multimodal feedback, while agency relied on the presence of visuomotor feedback, regardless of delay. However, in terms of a task, one would not rely on decreasing the amount of information given to the participant, nor on providing asynchronous visuomotor feedback, as these could hamper performance regardless of ownership or agency. A more appropriate factor for investigating the relation between these phenomena is then connectedness of a virtual hand. Tieri et al. (2015) found that participants only experienced ownership and vicarious agency (i.e. virtual arm moved but participants stayed still) when the arm was completely connected, and not when arm segments were missing. The main experiment from Chapter 4 showed that a disconnected hand also elicited ownership, but only when the hand was a mediated 3D projection of the own hand. It is possible that a virtual hand that is more realistic in appearance than the one used in the first pilot in Chapter 4 may show the expected ownership trend with respect to connectedness. Therefore, in this experiment, I hypothesize that ownership of the virtual hand will be higher in a connected condition than in a disconnected condition (*H2a*), and agency will similarly be higher in a connected condition than in a disconnected condition (*H2b*). I also hypothesize that as in the experiment in Chapter 4, there will be a shift in experienced hand-location (*H2c*) .

5.2 Agency and Performance

Wen et al. (2015) showed that action-feedback association (i.e. congruence between predicted and actual sensory information) and goal-directed inference (i.e. how well one was performing) both influenced the judgement of subjective agency in a continuous action task. They also showed that when the comparison between continuous action and feedback is difficult, then goal-directed inference plays a dominant role in judging agency. The experiment consisted of a key pressing task, where participants had to move a dot to a target by pressing arrow keys. They used conditions self-control versus assisted (i.e. incorrect key presses resulted in no movement, thus by definition better performance), and action delay of 100ms, 400ms and 700 ms. Performance was measured through duration, number of keys pressed and frequency of keys pressed. Agency ratings and the three performance measures all showed the same effects, namely they increased as delay increased, and were higher in the assisted condition than in the self-control condition. A multivariate analysis was used to estimate the relative influence of task performance on the sense of agency, and found assistance influenced the sense of agency indirectly via task performance, and delay influenced the sense of agency directly. The participants felt strong sense of agency when their task performance improved via computer assistance, even though a large proportion of their commands were not executed. This would suggest a correlation between agency and performance, even though the performance was not necessarily increased by the participants themselves.

Informally, one may suggest that a greater sense of being in control also coincides with better motor control. Possibly, experiencing more agency over a virtual hand makes the interaction performed by that hand feel more ‘natural’ to the participant than if there were no sense of agency. One could then suggest that the interaction may require fewer cognitive resources. The eliciting of a sense of agency is typically described to arise from a comparison between prediction and result, and Hon et al. (2013) showed that these comparisons are consciously performed. In their experiment, participants rated agency after moving a dot on a screen by pressing arrow keys, while they concurrently were asked to memorize two or six consonants which they were tested on, as a means of low and high load conditions, respectively. They found that agency ratings were significantly lower in the high load condition than in the low load condition. This would suggest that, since resources from a cognitive resource pool are already allocated in order to elicit the sense of agency, fewer resources remain for task execution, which is in contrast to the idea of performing better when the interaction is more natural. However, this does not explain how studies in VR consistently find higher task performance coinciding with a higher sense of agency (Egeberg et al., 2016; Laha et al., 2016). These studies are further discussed in the next section.

5.3 Ownership and Performance

Very few studies have examined the relation between body ownership and task performance directly. Older works have studied the relationship between performance and spatial presence. Snow (1998) confirmed that there was a positive relation between presence and performance of simple tasks related to a VR system's parameters, but that this relation was weak, and does not speak to the cause of this relation. Indeed, Welch (1999) describes the idea of presence causing better performance as a scientific urban legend, without there being evidence to support the causality.

In a more recent study on virtual wings in VR Egeberg et al. (2016), the authors investigated the role of different types of sensory feedback on body and wing ownership and agency using three conditions: only visuoproprioceptive feedback (no movement), only visuomotor feedback (rotating shoulders made the wings flap), and visuomotor and visuotactile feedback (during flapping). While visuoproprioceptive feedback alone did not in fact elicit any ownership or agency over the body or wings, the other two conditions did, where visuotactile feedback enhanced ownership and agency over the wings, but not the body. In a subsequent task participants were instructed to hit green balls and avoid red balls that were shot at them from a cannon, with or without visuotactile feedback. Although participants were equally well at hitting green balls in both visuomotor conditions, participants were able to avoid more red balls in the condition without visuotactile feedback. In summary, although participants experienced ownership and agency over both body and wings in both visuomotor conditions, performance was higher when there was no visuotactile feedback, which coincided with lower (but still positive) ownership and agency over the wings. This would suggest that a greater sense of ownership and/or agency can correspond to worse performance in a task, but the study lacks a correlation analysis, which makes it difficult to interpret whether such decrease in performance is caused by an increase in ownership alone, agency alone, or both, or neither of them.

Similarly, Laha et al. (2016) investigated the influence of control schema for a three-armed body in VR on task performance and, among other measures, body ownership and agency. They compared unimanual control (one wrist uses vertical and horizontal rotation for vertical and horizontal translation), bimanual (one wrist uses vertical, other horizontal) and head control (head uses vertical and horizontal rotation) of a third elongated arm protruding from the chest. Participants were instructed to touch three target cubes, each target being located in its own 3-by-3 array of cubes: to the left and 0.8m in front of the participant, centered and 1.3m in front of the participant, and to the right and 0.8m in front of the participant. The results showed that participants completed touching three cubes fastest in the head control condition followed by the uni- and bimanual conditions. Body ownership was higher in the unimanual and head conditions than in the bimanual condition, and agency was higher in the head condition than in the bimanual condition. In summary, the sense of ownership and agency coincided with greater performance, but again since the focus of the study was the control schemes rather than the relation between the three phenomena, no correlation analysis was performed.

In another recent study, Burin et al. (2019) examined the effects of ownership and agency on the ability to draw straight lines in VR. Viewpoint was altered (1PP using right hand versus 3PP using left hand), and ownership and agency of the virtual hand were measured after a baseline phase where participants were instructed to draw straight lines and simultaneously watched straight lines being drawn (matching their own drawing), and after a deviation phase where they instead saw curved lines being drawn (not matching their attempted straight lines). Correlation results showed that participants that reported a greater sense of body ownership, regardless of after which phase, were more inclined to follow the curved lines in their real drawings. That is, body ownership influenced motor actions. Moreover, there was a positive correlation between ownership of the virtual hand reported after the deviation phase and curve in the drawing, and also between agency after the deviation phase and curve in the drawing. It should be noted that ownership was maintained through the entire experiment in the 1PP, but agency was only experienced prior to the deviation phase. If one interprets a more curved drawing as a worse performance, since they were instructed to draw a straight line, than these results would suggest that when there is visuomotor discrepancy, greater experience of ownership over a virtual hand can result in worse performance. The authors explain that motor control can behave differently depending on whether the errors between predicted and actual feedback are causally attributed to the body or the environment.

Grechuta et al. (2017) drew evidence from brain activity studies that showed an overlap in the brain areas corresponding to body ownership, and those corresponding to motor control, and further investigated this overlap by means of a sensorimotor task in an RHI paradigm. Participants were instructed to press a button as soon as the fake hand was stroked under congruent visuotactile stimulation, incongruent haptic and incongruent visual stimulation. The authors found that ownership was higher and reaction times were lower in the congruent condition compared to the incongruent conditions, and a significant negative correlation between ownership and reaction time. It is explained that this confirms a functional role of ownership in the domain of motor control. In a next study the authors further examine this relationship, by arguing that ownership in RHIs using movement relies on an internal forward model, which in turn integrate signals from both proximodistal and purely distal sensory cues relevant to the task (Grechuta et al., 2019). Therefore, incongruent distal cues should impede both performance and the eliciting of ownership. This was confirmed in a VR air hockey experiment, where a condition with congruent distal cues was compared to a condition with incongruent distal cues. Both performance and sense of ownership were higher in the congruent condition than the incongruent conditions, while agency ratings did not differ between conditions but were nonetheless high, as was expected since there was no change in visuomotor congruency.

Lastly, in a very recent study the so far positive relationship between ownership and performance was challenged by the notion of risk of danger in the context of VR-based machinery teleoperation (Shin et al., 2021). Participants performed pick and place tasks on a conveyor belt, during which a ‘raw material’ had to be placed in a metal press machine for quick pressing (high risk) or slow pressing (low risk), using either a realistic hand or a robot hand. The results showed that body ownership significantly increased the risk perception during the operation, and was not moderated by actual risk of danger. Moreover, risk perception was negatively associated with work performance.

In summary, many studies have consistently found high ownership coinciding with high performance, but recent studies have found scenarios in which this suggested positive correlation becomes a negative correlation. However, one such scenario where intended sudden movement error was introduced seems unlikely in a scenario where high performance is desirable. Furthermore, although not supported through reasoning of allocated cognitive resources, studies repeatedly find coinciding agency and task performance. I hypothesize, therefore, that in a completely safe task performed in an ARSHI , any link between ownership and performance is mediated by agency (*H1b*).

5.4 Method

5.4.1 Design

To investigate the relation between embodiment and performance, this experiment included one factor ‘connectedness’, referring to the connectedness of the virtual hand, hereafter *Arm* and *Hand*, and was performed in a within-subjects design. I emphasize that the purpose of using the two conditions is to introduce variation in the responses to ownership and agency, in order to correlate both negative and positive ownership and agency responses to the performance values. The sense of ownership is measured by means of a questionnaire accompanied by skin conductance response (SCR) to a threat. The senses of agency and hand-location are also measured by means of a questionnaire. The Institutional Review Board of TNO positively recommended the execution of this experiment.

5.4.2 Participants

23 participants took part in the experiment, with mean age 30.5 (s.d. 9.5, range 19-47) of which 13 female and 10 male. Inclusion criteria were: between 18-50 years of age, right-handed, light skin color (to match as much as possible with the virtual arm model), not right arm/hand/finger amputee, no prosthetic on right arm/hand/finger, no scars or tattoos on right hand, and no experience with severe motion sickness or cybersickness.

5.4.3 Material

To create a video see-through AR setup, a ZED mini camera was mounted on to an HTC Vive. The ZED mini lens has a maximum field of view of 90°(horizontal) × 60°(vertical) × 100°(depth), and can reach 60 frames per second with a side-by-side output resolution of 2560×720 pixels. The HTC Vive offers a 110°field of view, a maximum refresh rate of 90 frames per second and a combined resolution of 2160×1200 pixels (1080×1200 pixels per eye). A Vive Tracker was placed on the table to determine the center of the interaction space, and another was strapped to the right wrist of the participant. The experiment was run on a Lenovo Legion T730-28ICO 90JF with a GEFORCE RTX 2080 Super graphics card and an Intel Core i9 processor. The project was created in Unity 2019.2.17f1 and Visual Studio 2019. The scene was visualized using SteamVR 1.15.19 and the SteamVR Unity Plugin 2.6.1. The 'VR Hands and FP Arms Pack' by NatureManufacture was used for the arm and hand, where in the latter case the arm was removed to create a single hand, see Figure 5.1. The 'Final IK' package by Rootmotion was used to allow the arm segments to move naturally. The 'Modern Combat Knife' by Float3D was used for the knife threat. A Biosemi set was used to measure SCR. The acquisition software ran on a separate Dell Latitude E6540 laptop. Using a Biosemi trigger interface, triggers were sent from the Unity project to the acquisition device through a serial port. The individual output measures were primarily analyzed using IBM SPSS Statistics 24 and supplemented by correlations analyses performed in RStudio 1.2.1335. Post hoc power analyses for the condition comparisons were performed in GPower 1.3, and for the correlations in IBM SPSS Statistics 27.

5.4.4 Procedure

Participants signed a consent form and washed their hands with a mild non-abrasive soap upon arrival at the laboratory. The experimenter attached four electrodes to the left hand: two to the thenar and hypothenar eminences, and two to the distal phalanges of the index and middle fingers. The experimenter helped the participant put on the HMD and the tracker on the right wrist. The participant then sat in an indicated start position at a table, with hands 30cm away from the table's edge and 50cm apart, while looking straight forward. The experimenter started the first condition, and the participant could see the room through the HMD, but with an added virtual arm or hand, see Figure 5.1. The practice session then started, during which the participant could move the virtual limb for 90 seconds to learn how its movement corresponded to the movement of the tracker. The position of the tracker determined the position of the virtual fingertip in both conditions, not the virtual wrist; the participants were not told to hold their hand in a specific shape. In the *Hand* condition, the virtual hand had fixed orientation. In the *Arm* condition, the virtual hand would rotate according to inverse kinematics.

After this, a first half ring would appear at 67cm from the table's edge and 65cm above the Tracker on the table, see Figure 5.1. The apparent full ring (torus) was 30cm (horizontal) × 30cm (vertical) × 6.5cm (depth). The participant was

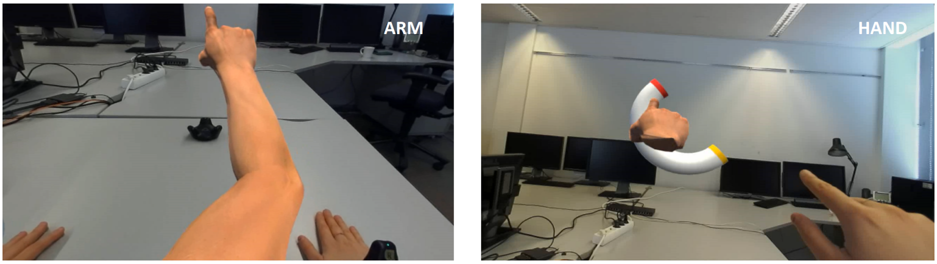


Figure 5.1: The participant's view during condition **(left)** *Arm* and **(right)** *Hand*. In the left image, the participant is in the practice session, in the right image the participant is completing a trial. The participant's goal during the trial was to trace the half ring from green (which turned yellow upon touch) to red.

instructed to touch the center of the intersection of the torus pipe, starting from the green side (0.75cm high, overlapping the end of the half ring), which would turn yellow once touched, to the red side (0.75cm high, overlapping the end of the half ring). Once the red side was touched, the half ring would disappear and a new ring would appear at a new random horizontal location, with the same vertical and depth location. Here, location refers to the center point of the full ring. The half ring would also rotate randomly in multiples of 45 degrees, and the drawing direction would also switch randomly from clockwise to counterclockwise.

After finishing 20 trials, the participant was asked to place their hand back in starting position, after which the virtual threat was launched: a knife would approach the virtual hand from the right and stop just before contact. After this the scene was turned off and the experimenter would orally ask seven questions to the participant in random order, who would answer on a scale from -3 to 3, representing "strongly disagree" to "strongly agree", respectively, see Table 5.1. The questions were the same as those used in the main experiment of Chapter 4, but the control questions were omitted. The participant was also given the opportunity to orally provide comments to the session, which were denoted by the experimenter. When the comment was lengthily or ambiguous, the experimenter would confirm the written piece with the participant. After this, the session was repeated using the other limb version, where ordering was counterbalanced across participants. After completing both sessions the participant was asked which hand they felt was more pleasant in use and were allowed to provide further comments about their decision.

Table 5.1: The list of questions that were asked after each condition, where the questions were presented in a random order, and answered on a 7-point Likert scale from "strongly disagree" to "strongly agree". Ownership is measured through three main questions that each reflect a different aspect of ownership: attribution, changing the body image, and the source of experienced sensations (Kilteni et al., 2012a; Guterstam et al., 2011), and the location questions cover three possible directions of change: multiple locations, full shift, and partial shift.

Q	*Question*	*Measure*
O1	It seemed as if the virtual hand was my hand.	Ownership
O2	It seemed as if I had three hands.	Ownership
O3	It seemed as if the movements I felt were caused by the movements of the virtual hand.	Ownership
A1	It seemed as if I was controlling the virtual hand.	Agency
L1	It seemed as if my left hand was at two different locations.	Location
L2	It seemed as if I felt the movements at the location of the virtual hand.	Location
L3	It seemed as if I felt the movements somewhere between the real and virtual hand.	Location

5.4.5 Analysis

For the ring tracing analysis, I wrote an algorithm in C# using Visual Studio 2015 to calculate the root mean square (RMS) deviation. Each half ring was divided into 180 bins, where each bin was a rectangular prism with frontal width equating to 1 degree of the half ring. The depth and frontal height of the prisms were chosen to be 14cm. Then all virtual finger tip data were sorted into these bins, and the smallest distance from the center of each prism to the sorted points was saved as the error ϵ for that bin. For empty bins, the error was automatically 7cm. The final performance measure was then equal to $1/\sqrt{\sum_i \epsilon^2/20}$. Using this inverse measure, a higher value indicates less deviation and thus better performance.

For the SCR analysis, I calculated each threat response by subtracting the average signal of the 5 seconds before the threat from the maximum signal in the 10 seconds after the threat. Here, 'threat' means the moment the virtual knife reached the proximity of the virtual hand. These responses were then transformed to a logarithmic scale: $\log_{10}(\Delta value + 1)$.

For the participants' comments, all individual statements were grouped into separate categories. Here, a statement means a meaningful expression concerning a single theme, and a comment could consist of multiple statements. Duplicate statements were removed, for example “it was difficult to move” in the first condition and “it was easier to move” in the second condition are counted as a single statement that only occurred once. The participants were not obliged to give a comment, nor were they given a maximum number of allowed comments, so the number of comments differs per person.

5.5 Results

5.5.1 Questionnaire Responses

For ease of reading, I discuss the questionnaire results by coding the Likert-responses to $-3, ..., 3$ corresponding with “strongly disagree”, ... , “strongly agree”, see Figure 5.2. Wilcoxon Signed Ranks Tests were performed on each pair of questionnaire responses. For agency (*A1*), there was a statistically significant difference between conditions, $Z = -3.281, p = 0.0002$ one-tailed, $1 - \beta = 0.996$; however, the agency ratings were in fact higher for the *Hand* condition (median 2) than the *Arm* condition (median 1), in contrast to the hypothesis. For shift in hand-location (*L2* and *L3*), both tests resulted in statistically significant differences between the conditions, $Z = -2.200, p = 0.014$ one-tailed and $Z = -1.841, p = 0.036$ one-tailed, respectively, but with low power, $1-\beta = 0.735$ and $1 - \beta = 0.615$. Participants rated a higher degree of both full and partial shifts in the *Hand* condition (medians 0 and 1, respectively) than in the *Arm* condition (medians -1 and 0, respectively). All other responses (*O1*, *O2*, *O3*, *L1*) did not differ significantly between conditions.

One-sample Wilcoxon Signed Rank Tests showed that the responses to *O1* in the *Arm* condition were significantly less than 0 ($Z = 56.50, p = 0.021, N = 23, 1 - \beta = 0.642$, two-tailed), with median -1. The responses to *A1* in both conditions were significantly greater than 0, with median 2 for condition *Hand* ($Z = 263.00, p = 0.0001, N = 23, 1 - \beta = 1.000$, two-tailed) and with median 1 for condition *Arm* ($Z = 168.00, p = 0.059, N = 23, 1 - \beta = 0.495$, two-tailed. Responses to *L2* in the *Arm* condition were significantly lower than 0, median -1, $Z = 48.00, p = 0.029, N = 23, 1 - \beta = 0.510$, two-tailed, and to *L3* in the *Hand* condition were significantly greater than 0, median 1, $Z = 170.00, p = 0.012, N = 23, 1 - \beta = 0.714$, two-tailed. All other responses (to *O1* for *Hand*, *O2*, *O3*, *L1*, *L2* for *Hand*, *L3* for *Arm*) did not significantly differ from 0.

5.5.2 Task Performance

After confirming that the performance values were normally distributed with a Shapiro-Wilk test ($W(23) = 0.960, p = 0.472$ for *Arm*, $W(23) = 0.952, p = 0.322$) for *Hand*), a paired samples t-test was used to compare performance between conditions. This showed that participants had statistically significantly higher deviation scores in the *Hand* condition (mean 2.208) than in the *Arm* condition (mean 1.956), $t(22) = -3.460, p = 0.002$, one-tailed, $1 - \beta = 0.973$. This means that participants performed better in the *Hand* condition than in the *Arm* condition.

5.5.3 SCRs

In two cases the SCR was not recorded due to equipment failure, thus data of the two relevant participants were excluded. No participants were classified as non-responders, that is, all participants demonstrated a difference in SCR within a single condition of more than 0.05μSiemens (Venables and Christie, 1980). The remaining data was confirmed to be normally distributed by Shapiro-Wilk tests, $W(21) = 0.957, p = 0.456$ for *Hand* and $W(21) = 0.948, p = 0.307$ for *Arm*. According to a paired-samples t-test, the SCRs did not differ significantly between conditions, $t(20) = -0.283, p = 0.780$ two-tailed. The SCRs were further analyzed for a habituation effect. Again, the remaining data was confirmed to be normally distributed, $W(21) = 0.963, p = 0.581$ for *First* and $W(21) = 0.904, p = 0.136$ for *Second*. A paired-samples t-test showed that there was no significant difference between the first and second condition, $t(20) = 1.587, p = 0.128$ two-tailed.

5.5.4 Correlation Ownership, Agency, Self-location and Performance

A repeated measures Spearman correlation from the R package rmcorr (Bakdash and Marusich, 2017) was performed on the following pairs:

- ownership measures amongst each other (*O1*-*O2*, *O1*-*O3*, *O1*-SCR, *O2*-*O3*, *O2*-SCR, *O3*-SCR)
- self-location measures amongst each other (*L1*-*L2*, *L1*-*L3*, *L2*-*L3*)
- ownership and agency (*O1*-*A1*, *O2*-*A1*, *O3*-*A1*, SCR-*A1*)
- ownership and performance (*O1*-Perf., *O2*-Perf., *O3*-Perf., SCR-Perf)

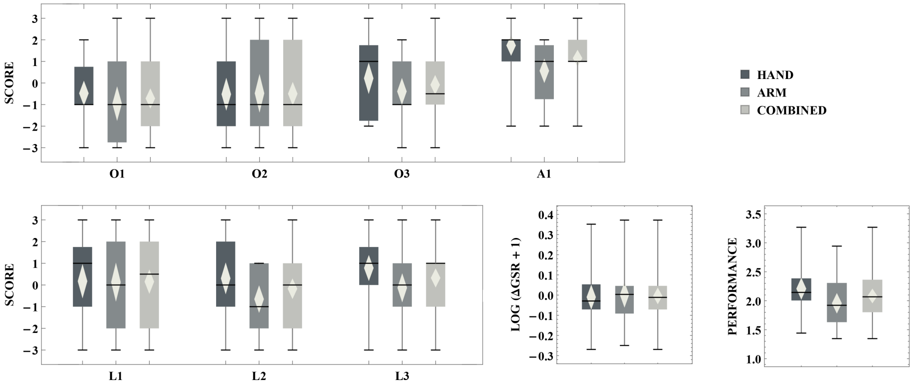


Figure 5.2: Boxplots of the results of **(top)** ownership and agency, **(bottom)** self-location, SCRs and performance scores. The diamonds are mean markers showing mean and mean standard error. For the questionnaire responses, the numbers -3, . . . , 3 correspond to responses "strongly disagree" to "strongly agree."

- agency and performance (*A1*-Perf.)

See section 5.6.4 for a brief discussion on the statistical approach. Here, I only report the significant correlations, which are visualized in Figure 5.3. Regarding correlations among the ownership measures, *O1*-*O2* was significantly correlated, $r = 0.355, p = 0.044, N = 23, 1 - \beta = 0.521$, and *O1*-*O3* and *O2*-SCR were statistically marginally correlated, $r = 0.342, p = 0.051, N = 23, 1 - \beta = 0.492$ and $r = -0.382, p = 0.080, N = 21, 1 - \beta = 0.542$, respectively. Note that *O2*-SCR is in the opposite direction than hypothesized. For correlations among the self-location measures, *L1*-*L2* and *L2*-*L3* were significantly correlated, $r = 0.429, p = 0.036, N = 23, 1 - \beta = 0.675$ and $r = 0.642, p = 0.0007, N = 23, 1 - \beta = 0.967$, respectively. For correlations between ownership and agency, *O1*-*A1* and *O3*-*A1* were statistically significantly correlated, $r = 0.424, p = 0.020, N = 23, 1 - \beta = 0.665$ and $r = 0.434, p = 0.017, N = 23, 1 - \beta = 0.685$, respectively. None of the ownership-performance pairs were significantly correlated. Lastly, regarding agency and performance, *A1*-Perf. was statistically significantly correlated, $r = 0.459, p = 0.024, N = 23, 1 - \beta = 0.734$. I remark that since the power is calculated assuming no repeated measures and still only 21 or 23 participants (i.e. assuming less data points), the actual power may be higher than reported here. To illustrate this, if I correlate averages over conditions of *O1*, *O2*, *O3* and SCR (i.e. those measures that did not differ significantly between conditions), then again significant correlations are found, *O1*-*O3* with $r = 0.532, p = 0.004, N = 23, 1 - \beta = 0.857$, and *O2*-SCR with $r = 0.486, p = 0.013, N = 21, 1 - \beta = 0.747$,

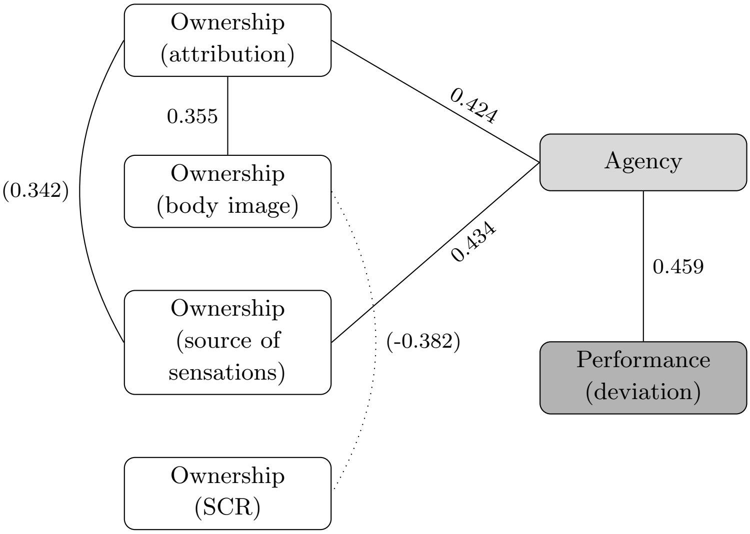


Figure 5.3: Map of all significant correlations with corresponding correlation coefficients for the repeated measures Spearman correlation on the data of both conditions together. The dotted line indicates a correlation direction opposite to hypothesized, and coefficients in brackets indicate marginal significance.

which have much higher power than the powers provided above. I therefore derive that the powers of correlation pairs *L1*-*L3*, *L2*-*L3*, *O1*-*A1*, *O3*-*A1* and *A1*-Perf. are presumably higher than 0.8.

5.5.5 Pleasantness and Comments

Three participants found the *Arm* condition more pleasant in use, and twenty participants found the *Hand* condition more pleasant in use. Regarding the participant comments, a full overview can be found in Table 5.2. The most frequently provided statements were: "movement in *Arm* was different than mine" and "movement in *Arm* was more difficult than in *Hand*".

Table 5.2: Participant comments, separated into individual statements and grouped into categories.

Category	*Statement*	*Frequency*
Hand rotation	unpleasant that hand in *Hand* did not rotate	2
	unpleasant that hand in *Arm* did not rotate	2
	hand in *Arm* had a different orientation than mine	2
	pleasant that hand in *Hand* did not rotate	1
Depth	difficult to see depth in *Hand*	3
	difficult to see depth in *Arm*	4
Embodiment	no ownership in *Hand*	2
	no ownership in *Arm*	3
	hand in *Hand* was a cursor	1
Movement	movement in *Arm* was different than mine	12
	movement in *Arm* was more difficult than in *Hand*	13
	movement in *Hand* was more difficult than in *Arm*	1
	movement in *Arm* was more sensitive	1
	pleasant that the hand in *Hand* had no hinges	2
Appearance	arm and hand in *Arm* had strange appearance	3
	hand in *Hand* had strange appearance	2
Visual presence	absence of arm in *Hand* did not matter	1
	presence of arm in *Arm* was pleasant	1
	arm occupied too much of my view	3
	arm occupying my view did not matter in regards to controlling it	1
	real right hand not always in view so did not feel like three hands	1
Experiment	it was tiring	2
	meta statement on setup or execution	7

5.6 Discussion

The goal of this study was to examine the relation between ownership and performance in the ARSHI. Participants were asked to trace a half ring as accurately as possible in two conditions: a connected *Arm* condition which was expected to result in high virtual hand ownership and agency ratings, and a disconnected *Hand* condition which was expected to result in low ratings. The results showed that ownership ratings did not differ between conditions, and, surprisingly, that agency ratings were higher in the *Hand* condition than in the *Arm* condition. In the correlation analyses there was a positive correlation between ownership and agency ratings, as well as between agency ratings and performance, but not between ownership ratings and performance.

5.6.1 Ownership

I had hypothesized that participants would experience greater ownership in the *Arm* condition than in the *Hand* condition based on other works (Tieri et al., 2015). For all ownership responses including SCRs, there were no significant differences between the two conditions, thus *H2a* is rejected.

There were weak correlations between different ownership measures but all with low power, making it difficult to draw conclusions. Since multiple RHI studies in reality and VR have found positive correlations between various qualitative and quantitative measures of ownership, I suspect that the experience in the ARSHI, i.e. extending the real body with a virtual limb, greatly differs from these other real and virtual RHIs, where the real limb/body is replaced with a fake limb/body. Indeed, the study by Feuchtner and Müller (2017), that studies ways to present the virtual hand in an RHI in AR in order to interact with real objects, shows similar low responses regarding direct attribution ("own hand", *O1*), in conditions "*abstract hand*" (similar to *Hand*) and "*arm without inpaint*" (similar to *Arm*). The responses to *O2* on body image ("three hands"), on the other hand, showed no significant difference between conditions, whereas the study by Feuchtner and Müller would suggest greater experiencing of three hands in the *Arm* condition than in the *Hand* condition.

Moreover, although the ratings were quite spread, the majority did not differ significantly from 0 and the responses to *O1* for *Arm* were significantly lower than 0, that is, the majority of participants did not experience any degree of direct attribution. In the following, I suggest three reasons for finding overall negative ownership results and thus also no difference between conditions.

First, one probable factor is the complex notion of visual realism in AR. Ownership studies in VR have similarly found mixed results on the effect of hand realism. For example, while Pyasik et al. found that using a 3D scan of the real hand in VR resulted in greater ownership over the virtual hand than over a virtual hand model, Jo et al. found the opposite effect, namely that a cartoon version of the participant with matching clothes elicited more body ownership than a 3D scan of the participant. In the current study, five statements were made by participants about the strange appearance of the virtual arm/hand (see Table 5.2), of which none explicitly referred to a mismatch between the appearance of their limb and the virtual limb. In the main experiment of Chapter 4, I indeed found that with a virtual projection of the real hand, participants still found it difficult to accept the virtual hand as their own. I discussed that in AR participants may have to be more willing to believe that what is fake is not fake, and I indeed found a positive relation between the participants' immersive tendency and their ownership responses.

Secondly, it may be that fakeness of the virtual hand was not only experienced in the visual aspect, but also its movement. The frequent comments about the movement of the virtual limb in the connected *Arm* condition may illustrate that there was a greater expectation about the abilities of the connected (i.e. more realistic) arm, in comparison to the disconnected (i.e. less realistic) hand. I emphasize that the positioning of the virtual fingertips used the tracked wrist data identically in both conditions, but the rotation of the wrist differed as a result of the use of inverse kinematics in the *Arm* condition and nothing in the *Hand* condition. This dip as a result of unfulfilled expectation resembles the popularly referenced uncanny valley effect, and what was found with respect to multidimensional realism in second pilot in Chapter 4. In the current study, the visible arm as opposed to no arm may have also become a form of distraction, since all questions only referred to the hand, reducing the visible arm to be experienced as 'noise'. Furthermore, this 'noise' is not related to the more simple notion of number of presented distracting pixels (i.e. in *Hand* no arm pixels and in *Arm* many arm pixels), since Okumura et al. found higher ownership ratings of a supernumerary virtual hand in AR in a condition with high arm opacity (i.e. more pixels) than in a condition with low arm opacity (i.e. less pixels). However, it is difficult to attribute these findings solely to the uncanny valley, as the evidence of whether it even exists is mixed.

Lastly, it is also possible that the relation between realism (in whatever form it may take) and embodiment in AR differs in nature that the corresponding relation in reality and VR. That is, it may not be straightforward to expect an a priori positive relation between embodiment and realism, because a third virtual arm may be more disturbing in AR than an abstract tool. The study by Tieri et al. found a typical ownership experience, but the setups differ fundamentally in that those participants did not actually move the virtual hand themselves, nor were they subjected to a performance related task. Because of this different context, I suggest that in the current study, even if users had not experienced the movement as improper for the 'more realistic' connected arm (e.g. by using different inverse kinematics), then possibly still the floating hand may not have been experienced as more unrealistic than a third arm, as would be suggested by the uncanny valley effect discussion above.

Together, these findings suggest that (1) visual realism is more complex in AR than in reality and VR and requires more willingness to believe in order to accept a virtual object as real, (2) increasing realism in a single dimension can cause expectations in other realism-dimensions, that, when not fulfilled, may lead to an overall less pleasant experience, and (3) the relation between embodiment and realism in AR fundamentally differs from the corresponding relation in reality and VR.

5.6.2 Agency and Self-location

I hypothesized that participants would experience more agency in the *Arm* condition than in the *Hand* condition, also based on the results by Tieri et al.. The results showed that in both conditions agency ratings were highly positive, but that the responses in the *Hand* condition were actually higher than in the *Arm* condition, thus *H2b* is rejected. Also, I hypothesized participants would experience a shift in hand-location in both conditions, and the results showed that this was only the case for 'partial shift' (*L3*) in the *Hand* condition, while responses to 'full shift' (*L2*) in the *Arm* condition were actually negative and all others approximately 0, thus *H2c* is rejected.

Regarding agency, in the previous section on ownership, I discussed how the setup of the study by Tieri et al. differed fundamentally from this setup in terms of the cause of movement and the experimental context. Due to the absence of participant movement, the authors only measure vicarious agency, that is, the feeling of being the agent of others' actions. I acknowledge it may have been to simplistic to assume a similar sense of body agency would occur in this study. From the comments, it became clear that the participants experienced some form of discrepancy in the movement of the virtual limb in the *Arm* condition, despite identical virtual fingertips positioning mechanisms in both conditions. I expect that this experience was largely caused by having to move with a specific purpose rather than just synchronous, but further meaningless, movement as is typical in a typical RHI. Participants tried their best, as instructed, to trace the half ring, a seemingly straightforward and simple task, but found that it was more difficult than expected before execution. This may have led to frustration and automatically thinking they were performing badly. In the case of the connected arm, they may have moreover been preoccupied by the way the arm segments were moving differently to theirs, leading to even more frustration, and in turn the feeling of performing worse. This would suggest, in line with the results of the study by Wen et al., that agency decreased as estimated performance decreased, even though they were provided no information regarding performance compared to the other participants or the other condition.

Regarding hand-location, I expected that successfully reaching an object in peripersonal space would result in a change in experienced hand-location, although it was uncertain whether this would take form in a shift or separation of normal hand-location. This was different than in the experiment from Chapter 4, where participants did not have to actively reach to the boundaries of their personal space. However, I saw that such a change did not occur with the exception of the partial shift experienced only in the *Hand* condition. To explain the results, I turn to the definition of self-location, namely the volume in space where one feels to be located, which in daily life coincides with body-space, meaning one feels self-located inside the physical body (Kilteni et al., 2012a). I found that participants made seven statements about having difficulty seeing depth, three of them occurring in the *Hand* condition and four in the *Arm* condition (see Table 5.2). Possibly, then, participants struggled to make a mental spatial model, and as a result, they could not reliably say whether the virtual hand felt located in the personal or peripersonal space, as suggested by the 0-level responses rather than negative responses.

5.6.3 The Relation between Ownership, Agency and Performance

I hypothesized that there would be a positive correlation between experienced ownership and performance of the virtual hand, and that this correlation would be mediated by agency. The results showed a significantly better performance in the *Hand* condition than the *Arm* condition. There was no significant correlation between ownership and performance, thus *H1a* is rejected. However, there were moderate positive correlations between ownership (in terms of direct attribution (*O1*) and the source of experienced sensations (*O3*)) and agency, and a moderate positive correlation between agency and performance, thus *H1b* is accepted. In the following I attempt to place these findings within existing embodiment frameworks in order to present a possible causality. Since the focus here is solely bottom-up and top-down processing mechanisms, I do not include the hypothesis testing level described in Chapter 4.

It has been demonstrated that the top-down processing mechanism depends on whether there is self-generated movement or not (Grechuta et al., 2017). When the self-movement is congruent, ownership is high, even in cases with incongruent body characteristics, indicating that when a participant actively moves, the processing mechanism no longer depends on the internal body model as in the traditional non-moving RHI, but rather on predictive forward models. In this study, this would mean that ownership would be high regardless of connectedness of the virtual hand. However, as explained in the previous section on ownership, there was an experience of incorrect movement by the virtual hand beyond the positioning of the fingertip. Translating to internal forward models, the internal prediction about how the hand as a whole would move did not match the actual movement, since only the position of the fingertip matched their own real wrist movements and the rest of the limb did not correspond to their own movements, thus the premise of congruent self-movement no longer holds, and the sense of ownership once again relies on the internal body model. As discussed in Section 5.6.1, the low ownership results may then have been the result of incongruence in visual appearance of the virtual hand.

However, although their was no overall experience of ownership in this study, ownership was found to be positively correlated to agency, but importantly, not to performance. This latter finding contrasts what was found by Grechuta et al. (2017). The authors explain that body ownership is a result of multimodal integration, and the results of this integration can be used to modulate performance, or, put in terms of Bayesian inference for decision making, congruent information reduces perceptual ambiguity which can enhance motor response. This simply suggests, however, a causality of multimodal integration (i.e. the creation of a mental model) to the eliciting of ownership, and of multimodal integration to motor control, but does not restrict the increasing of performance to a case where ownership is also increased. The results confirm this, since the connectedness had no effect on body ownership, but did affect performance. In other words, such a correlation can only exist when the factor used to alter experiences of ownership is related to motor control.

When the factor is not related to motor control, the results suggest that the sense of agency can still be altered, even if it does not alter the sense ownership. A possible explanation is that the participants did not actually experience body agency in both conditions, but external (tool) agency in one or both conditions. Kalckert and Ehrsson (2012) provide evidence that these are distinct experiences, where body agency may only be related to transfer of sensorimotor integration mechanisms to the virtual hand (as is body ownership), whereas external agency relies on sensory predictions based on actions and goals from learned experiences. However, the positive correlation between ownership and agency in the present study contradicts the notion that participants exclusively experienced external agency rather than body agency, thus it is unlikely that participants reported external agency instead of body agency. Furthermore, Kalckert and Ehrsson have suggested a directional causality between the sense of ownership and the sense of body agency, since there is a general tendency to ascribe agency to an owned body part, and found little support for the opposite causality.

The findings by Wen et al. (2015) would further suggest that the positive correlation between the agency ratings and the performance scores was due to a causality from performance to agency, however in their study, participants were clearly aware of being assisted with a resulting positive effect on performance. In the current study, participants were completely unaware of their real-time performance, nor were they provided any means by which they could deduce how well they performed overall, thus translating these causality findings to this study is not straightforward. The opposite causality would seem to contradict the evidence that creating a sense of agency in itself requires the allocation of cognitive resources (Hon et al., 2013). However, it is possible that the number of resources subsequently required in the sensorimotor task are lower, precisely due to the elicited sense of agency. If this gain outweighs the resources required for agency, then this would indeed suggest a causality from agency to performance.

To summarize, these results do not support a direct relation between the sense of ownership and sensorimotor task performance. Instead, there was evidence for a relation between the sense of ownership and the sense of agency, and also between the sense of agency and performance. The design of the experiment does not allow conclusions to be drawn regarding causality. However, based on the above discussion, there is support for the following causality: the sense of ownership over a third virtual hand in AR influences the sense of agency over that virtual hand, which in turn influences performance in a sensorimotor task performed by that hand. However, I emphasize that this was not the focus of this study, and further investigating possible causalities is an important topic for future research.

5.6.4 Limitation

In order to examine correlations between ownership, agency and performance, the formal statistical method would be to perform a repeated measures correlation on the data of both conditions together, taking into account that the questionnaire data is ordinal and the SCRs and performance values are both continuous. Unfortunately, such an analysis does not exist in SPSS or in R and writing an appropriate package is outside the scope of this study. Therefore the closest alternative, namely considering all data as continuous, was chosen. Although this may slightly affect the statistical results, this should not happen to such as extent that the conclusions are no longer valid. Alternatives for correlation analysis, such as a Spearman correlation on the combined data (i.e. not repeated measures) or two Spearman correlations on the split data per condition, are listed in Table 5.3. Note that the significant correlations found in these ways are not inconsistent with the results presented above. For completeness, a linear regression is not applicable to these results since both to be analysed measures are random variables and not fixed without error, and is therefore not performed in this study.

5.7 Conclusion

This chapter examined the practical benefits of the sense of embodiment. The objective of this experiment was to investigate whether the often suggested relation between the sense of ownership and task performance is established through the sense of agency in an ARSHI. In this regard, the results showed that there was no direct relation between the sense of ownership and sensorimotor task performance, but an indirect relation through the sense of agency. This finding has implications for serious domains in which optimal performance is crucial, such as the area of teleoperations.

Table 5.3: Alternative statistical approaches to the correlation analysis. *Combined* corresponds to the Spearman correlation on the combined data (not repeated measures), and *Hand* and *Arm* to the Spearman correlation on the data split by condition; in *Arm* the GSR-Perf. pair uses a Pearson correlation. Fields marked by a (-) indicate that the corresponding pair violated the monoticity assumption, and pairs where all three analyses were in violation are omitted. Statistically significant correlations (one-tailed) are marked by a (*) and marginally significant correlations by a (•).

Measure(s)	*Pair*	*Combined*	*Hand*	*Arm*
Ownership	Q1-Q3	$\rho = 0.449$ $p = 0.001$* $N = 46$	$\rho = 0.192$ $p = 0.190$ $N = 23$	$\rho = 0.627$ $p = 0.001$* $N = 23$
	Q1-GSR	$\rho = 0.024$ $p = 0.439$ $N = 42$	-	-
	Q2-GSR	$\rho = -0.411$ $p = 0.003$* $N = 42$	$\rho = -0.601$ $p = 0.002$* $N = 21$	$\rho = -0.279$ $p = 0.111$ $N = 21$
Own. - Age.	Q1-Q4	-	-	$\rho = 0.289$ $p = 0.091$ $N = 23$
	Q2-Q4	-	-	$\rho = 0.139$ $p = 0.264$ $N = 23$
	Q3-Q4	$\rho = 0.227$ $p = 0.065$• $N = 46$	-	$\rho = 0.363$ $p = 0.044$* $N = 23$
Own. - Perf.	GSR-Perf.	-	-	$r = -0.081$ $p = 0.363$ $N = 21$
Self-location	Q5-Q6	$\rho = 0.257$ $p = 0.042$* $N = 46$	-	$\rho = 0.634$ $p = 0.001$* $N = 23$
	Q5-Q7	$\rho = 0.518$ $p = 0.0001$* $N = 46$	-	

Chapter 6

Conclusion

In this dissertation, I examined supernumerary embodiment as a form of the player-avatar link in multimodal AR games. In Chapter 2 I discussed imprecise and overlapping definitions of real and virtual objects, and presented a redefinition of the elements that make up multimodal AR, namely real, mediated, and virtual stimuli. Since mediated stimuli can be part of what is augmented or the augmentation itself, I present the phrases basis and augmentation to signify the original environment and the new. Lastly, I present a classification scheme for AR that relies on how the multiple modalities are spread across basis and augmentation. Together, these steps create a clearer picture of what is and what is not AR. The new definitions give a clear description of the components that make up multimodal AR, and how they can be used to create certain experiences: a user will react differently to seeing a mediated arm than to seeing a virtual arm.

In Chapter 3, I present an overview of studies on player-avatar links from an interdisciplinary perspective. While the studies from Humanities typically focus on traditional digital games and varied widely in view, many studies from Natural Sciences focus on embodiment of virtual avatars in general. In both cases, the mediation and willingness to belief seemed to play an important role. When translating these views to AR games, I described specific avatar representations and basis-augmentation combinations for which it was not clear whether embodiment as a player-avatar link was feasible. Specifically, a non-overlapping avatar with simultaneously visible own body, and partially overlapping avatar required further examination, and it was suggested that a smaller gap between basis and augmentation may facilitate embodiment.

Chapter 4 describes four experiments where these avatar representations and different basis-augmentation combinations are examined in the experience of embodiment (ownership, agency and self-location). In the first pilot experiment, a partially overlapping avatar in the form of a third virtual hand with visuotactile and visuomotor stimulation along the real body did not elicit any experience of ownership. In the second pilot experiment, a non-overlapping avatar in the form of a second virtual body or body-shaped object with visuomotor stimulation along the mediated body also did not elicit an experience of ownership. In this case, there was also possibly a realism interaction effect at play between appearance and movement, resembling the uncanny valley effect. Specifically, a realistic appearance with comparably less realistic movement was seen as equally realistic as a less realistic appearance with comparably more realistic movement, even though both movements were in fact identical.

In the third main experiment, a mediated 3D projection of the own hand was shown along the mediated body with realistic visuotactile and visuomotor stimulation, and this resulted in positive ownership responses. In all three experiments there was also a strong experience of agency for congruent visuomotor stimulation, and generally no or weak changes in experienced hand-location. In a fourth experiment, I examined the effect of immersive tendency on the experience of ownership and found a positive correlation between reported immersive tendency and ownership ratings. I finally present a model for ownership in the ARSHI that incorporates media and user variables in two separate processing levels: one for model construction and one for hypothesis testing. By incorporating these variables, the model on the one hand justifies frequently occurring individual differences in media-based RHI experiments, and on the other hand further illustrates fundamental differences between experiencing embodiment in AR versus non-mediated situations.

In Chapter 5, I explore the purpose of experiencing embodiment in the context of human visuomotor behavior. Specifically, I inspect task performance improvement as a practical benefit to the experience of ownership in AR, in a task using a third virtual hand along the mediated body. Such a relation had been suggested in the literature for decades, but lacked both a formal premise and empirical evidence. I hypothesized that agency would play a mediating role in any possible correlation and I found that there was actually no direct correlation between ownership responses and performance, but there was indeed an indirect correlation through reported agency. I discuss that this correlation is likely a causality relation from ownership to agency to performance. Again, the ownership results showed there was a realism interaction effect between appearance and movement.

AR research is still in its infancy, not only technologically but also theoretically, and even more so in understanding the unique experiences it produces. The work presented in this dissertation takes steps in these directions. While the studies presented in this dissertation provide many areas for future research, I shall here only highlight those that are, in my opinion, the most critical.

Multimodality - While current AR games and applications are typically solely visually-based, I urge the research community to also accommodate multimodal experiences. The main experiment of Chapter 4 found that the experience of ownership was strongest in the conditions with multiple forms of multimodal stimulation, which would more generally suggest that stronger experiences can occur when catering to more senses in AR. It is clear that this step will likely be more complex than the already impressive area of multimodality in VR, but similarly, the gain in experience is expected to be clearly beneficial, be it experiencing new game worlds for fun, or improving performance in serious training purposes.

Realism and belief - Throughout the experiments presented in this dissertation, there seemed to be a repeated interaction effect between different forms of realism that may have caused similar (low) ownership ratings across conditions, namely that realism in aspect A could be diminished by comparably less realism in aspect B, and conversely that low realism in aspect A could be improved by comparably more realism in aspect B. While this seems to have been partially overcome by using a the shape and appearance of the own hand and realistic multimodal feedback, there was still a reservation among participants to accept the supernumerary limb as real. As suggested throughout this dissertation, it seems that belief may play a bigger role in AR embodiment studies than those in VR and reality. As is clear from VR studies, it is not always necessary to increase the realism in the form of quality of the stimulation to ensure acceptance, but can be aided by the quantity of included senses and level of interaction when presented accordingly to the expectation of the user. In this regard, analogous AR studies may be beneficial to future research and developers. Moreover, if belief proves to be crucial in AR, then studying personal characteristics linked to media use will as well.

With this thesis I disambiguated the components that make up and determine experiences in multimodal AR, studied how different avatar representations and basis-augmentation combinations affect embodiment in AR, presented a validated processing model for the experience of hand ownership in AR, and showed that there is an indirect relation between ownership of a supernumerary hand and task performance through agency.

Bibliography

(1993). Doom, id Software. [MS-DOS].

(2012). Luigi's Ghost Mansion, Nintendo Land. [Nintendo Wii U].

(2016). Fragments, Asobo Studio. [Hololens].

(2016). Holoportation, Microsoft Research. [Hololens].

(2016). Young conker, Asobo Studio. [Hololens].

Aarseth, E. (1997). *Cybertext: Perspectives on ergodic literature.* John Hopkins University Press, Baltimore, MD, USA.

Armel, C. and Ramachandran, V. (2003). Projecting sensations to external objects: evidence from skin conductance response. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1523):1499–1506.

Asai, T., Mao, Z., Sugimori, E., and Tanno, Y. (2011). Rubber hand illusion, empathy, and schizotypal experiences in terms of self-other representations. *Consciousness and Cognition*, 20(4):1744–1750.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B. (2001). Recent advances in augmented reality. *Computer Graphics and Applications, IEEE*, 21(6):34–47.

Bach-y Rita, P. and Kercel, S. W. (2003). Sensory substitution and the human–machine interface. *Trends in cognitive sciences*, 7(12):541–546.

Bakdash, J. Z. and Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in psychology*, 8:456.

Banks, J. (2015). Object, me, symbiote, other: A social typology of player-avatar relationships. *First Monday*, 20(2).

Bayart, B. and Kheddar, A. (2006). Haptic augmented reality taxonomy: haptic enhancing and enhanced haptics. In *Proceedings of EuroHaptics*, pages 641–644.

Bayliss, P. (2007). Beings in the game-world: characters, avatars, and players. In *Proceedings of the 4th Australasian conference on Interactive entertainment*, page 4. RMIT University.

Behm-Morawitz, E. (2013). Mirrored selves: The influence of self-presence in a virtual world on health, appearance, and well-being. *Computers in human Behavior*, 29(1):119–128.

Biocca, F. (1999). The cyborg's dilemma: Progressive embodiment in virtual environments. *Human Factors in Information Technology*, 13:113–144.

Bommel, J. v. (2017). Presence and embodiment in augmented reality. Master's thesis, Utrecht University.

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

Braun, N., Debener, S., Spychala, N., Bongartz, E., Sörös, P., Müller, H. H., and Philipsen, A. (2018). The senses of agency and ownership: a review. *Frontiers in psychology*, 9:535.

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

Burin, D., Pyasik, M., Salatino, A., and Pia, L. (2017). That's my hand! Therefore, that's my willed action: How body ownership acts upon conscious awareness of willed actions. *Cognition*, 166:164–173.

Chen, W.-Y., Huang, H.-C., Lee, Y.-T., and Liang, C. (2018). Body ownership and the four-hand illusion. *Scientific reports*, 8(1):2153.

Crick, T. (2010). The game body: Toward a phenomenology of contemporary video gaming. *Games and Culture*, 6(3):259–269.

Egeberg, M. C., Lind, S. L., Serubugo, S., Skantarova, D., and Kraus, M. (2016). Extending the human body in virtual reality: Effect of sensory feedback on agency and ownership of virtual wings. In *Proceedings of the 2016 Virtual Reality International Conference*, pages 1–4.

Ehrsson, H. H. (2009). How many arms make a pair? perceptual illusion of having an additional limb. *Perception*, 38(2):310–312.

Farrow, R. and Iacovides, I. (2014). Gaming and the limits of digital embodiment. *Philosophy & Technology*, 27(2):221–233.

Feuchtner, T. and Müller, J. (2017). Extending the body for interaction with reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 5145–5157. ACM.

Fischer, J., Bartz, D., and Straber, W. (2005). Stylized augmented reality for improved immersion. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pages 195–202. IEEE.

Folegatti, A., Farnè, A., Salemme, R., and de Vignemont, F. (2012). The rubber hand illusion: Two's a company, but three's a crowd. *Consciousness and Cognition*, 21(2):799–812.

Germine, L., Benson, T. L., Cohen, F., and Hooker, C. I. (2013). Psychosis-proneness and the rubber hand illusion of body ownership. *Psychiatry Research*, 207(1):45–52.

Gilbers, C. (2017). The sense of embodiment in augmented reality: A third hand illusion. Master's thesis, Utrecht University.

Grechuta, K., Guga, J., Maffei, G., Rubio Ballester, B., and Verschure, P. F. (2017). Visuotactile integration modulates motor performance in a perceptual decision-making task. *Scientific reports*, 7(1):1–13.

Grechuta, K., Ulysse, L., Rubio Ballester, B., and Verschure, P. F. (2019). Self beyond the body: action-driven and task-relevant purely distal cues modulate performance and body ownership. *Frontiers in Human Neuroscience*, 13:91.

Guterstam, A. and Ehrsson, H. H. (2012). Disowning one's seen real body during an out-of-body illusion. *Consciousness and cognition*, 21(2):1037–1042.

Guterstam, A., Petkova, V., and Ehrsson, H. (2011). The illusion of owning a third arm. *PloS one*, 6(2):e17208.

Hon, N., Poh, J.-H., and Soon, C.-S. (2013). Preoccupied minds feel less control: Sense of agency is modulated by cognitive load. *Consciousness and cognition*, 22(2):556–561.

Hou, J., Nam, Y., Peng, W., and Lee, K. M. (2012). Effects of screen size, viewing angle, and players' immersion tendencies on game experience. *Computers in Human Behavior*, 28(2):617–623.

IJsselsteijn, W. A., de Kort, Y. A. W., and Haans, A. (2006). Is this my hand i see before me? the rubber hand illusion in reality, virtual reality, and mixed reality. *Presence: Teleoperators and Virtual Environments*, 15(4):455–464.

Jeon, S. and Choi, S. (2009). Haptic augmented reality: Taxonomy and an example of stiffness modulation. *Presence: Teleoperators and Virtual Environments*, 18(5):387–408.

Jin, S. and Park, N. (2009). Parasocial interaction with my avatar: Effects of interdependent self-construal and the mediating role of self-presence in an avatar-based console game, Wii. *CyberPsychology & Behavior*, 12(6):723–727.

Jo, D., Kim, K., Welch, G. F., Jeon, W., Kim, Y., Kim, K.-H., and Kim, G. J. (2017). The impact of avatar-owner visual similarity on body ownership in immersive virtual reality. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pages 1–2.

Jung, S., Bruder, G., Wisniewski, P. J., Sandor, C., and Hughes, C. E. (2018). Over my hand: Using a personalized hand in vr to improve object size estimation, body ownership, and presence. In *Proceedings of the Symposium on Spatial User Interaction*, pages 60–68. ACM.

Kalawsky, R., Stedmon, A., Hill, K., and Cook, C. (2000). A taxonomy of technology: Defining augmented reality. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 44, pages 507–510. SAGE Publications.

Kalckert, A. and Ehrsson, H. H. (2012). Moving a rubber hand that feels like your own: a dissociation of ownership and agency. *Frontiers in human neuroscience*, 6:40.

Kállai, J., Hegedüs, G., Feldmann, Á., Rózsa, S., Darnai, G., Herold, R., Dorn, K., Kincses, P., Csathó, Á., and Szolcsányi, T. (2015). Temperament and psychopathological syndromes specific susceptibility for rubber hand illusion. *Psychiatry Research*, 229(1):410–419.

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

Kilteni, K., Normand, J.-M., Sanchez-Vives, M., and Slater, M. (2012b). Extending body space in immersive virtual reality: a very long arm illusion. *PloS one*, 7(7):e40867.

Klevjer, R. (2006). *What is the avatar? Fiction and embodiment in avatar-based singleplayer computer games.* PhD thesis, University of Bergen, Norway.

Klevjer, R. (2012). Enter the avatar: The phenomenology of prosthetic telepresence in computer games. In *The philosophy of computer games*, pages 17–38. Springer.

Laha, B., Bailenson, J. N., Won, A. S., and Bailey, J. O. (2016). Evaluating control schemes for the third arm of an avatar. *Presence: Teleoperators and Virtual Environments*, 25(2):129–147.

Lee, K. M. (2004). Presence, explicated. *Communication Theory*, 14(1):27–50.

Lindeman, R. W. and Noma, H. (2007). A classification scheme for multi-sensory augmented reality. In *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, pages 175–178. ACM.

Linderoth, J. (2005). Animated game pieces. avatars as roles, tools and props. In *Proceedings of the Aesthetics of Play Conference*, pages 14–15, Norway. University of Bergen.

Lombard, M. and Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2):JCMC321.

Longo, M., Schüür, F., Kammers, M., Tsakiris, M., and Haggard, P. (2008). What is embodiment? A psychometric approach. *Cognition*, 107(3):978–998.

Lugrin, J.-L., Latt, J., and Latoschik, M. (2015). Anthropomorphism and illusion of virtual body ownership. In *Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments*, ICAT - EGVE '15, pages 1–8, Aire-la-Ville, Switzerland. Eurographics Association.

Mann, S. (2002). Mediated reality with implementations for everyday life. *Presence Connect*.

Maselli, A. and Slater, M. (2013). The building blocks of the full body ownership illusion. *Frontiers in human neuroscience*, 7:83.

Maselli, A. and Slater, M. (2014). Sliding perspectives: dissociating ownership from self-location during full body illusions in virtual reality. *Frontiers in human neuroscience*, 8:693.

McDonnell, R., Breidt, M., and Bülthoff, H. H. (2012). Render me real? investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics (TOG)*, 31(4):1–11.

Merleau-Ponty, M. (1945). *Phénomènologie de la perception.* Gallimard, Paris, France.

Milgram, P. and Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, 77(12):1321–1329.

Mori, M., MacDorman, K. F., and Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2):98–100.

Müller, T. (2015). Towards a framework for information presentation in augmented reality for the support of procedural tasks. In *International Conference on Augmented and Virtual Reality*, pages 490–497. Springer.

Nash, E., Edwards, G., Thompson, J., and Barfield, W. (2000). A review of presence and performance in virtual environments. *International Journal of human-computer Interaction*, 12(1):1–41.

Newman, J. (2002). The myth of the ergodic videogame. *The International Journal of Computer Game Research*, 2(1):1–17.

Newport, R., Pearce, R., and Preston, C. (2010). Fake hands in action: Embodiment and control of supernumerary limbs. *Experimental Brain Research*, 204(3):385–395.

Okumura, K., Ora, H., and Miyake, Y. (2020). Investigating the hand ownership illusion with two views merged in. *Frontiers in Robotics and AI*, 7.

Perez-Marcos, D., Sanchez-Vives, M., and Slater, M. (2012). Is my hand connected to my body? The impact of body continuity and arm alignment on the virtual hand illusion. *Cognitive neurodynamics*, 6(4):295–305.

Petkova, V. and Ehrsson, H. (2008). If I were you: perceptual illusion of body swapping. *PloS one*, 3(12):e3832.

Punpongsanon, P., Iwai, D., and Sato, K. (2015). Softar: Visually manipulating haptic softness perception in spatial augmented reality. *IEEE transactions on visualization and computer graphics*, 21(11):1279–1288.

Pyasik, M., Burin, D., and Pia, L. (2018). On the relation between body ownership and sense of agency: A link at the level of sensory-related signals. *Acta Psychologica*, 185:219–228.

Pyasik, M., Tieri, G., and Pia, L. (2020). Visual appearance of the virtual hand affects embodiment in the virtual hand illusion. *Scientific reports*, 10(1):1–11.

Ratan, R. A. (2012). Self-presence, explicated: Body, emotion, and identity extension into the virtual self. *Handbook of Research on Technoself: Identity in a Technological Society*, page 322.

Ratan, R. A. and Dawson, M. (2016). When mii is me: A psychophysiological examination of avatar self-relevance. *Communication Research*, 43(8):1065–1093.

Ryan, M.-L. (1999). Immersion vs. interactivity: Virtual reality and literary theory. *SubStance*, 28(2):110–137.

Schaefer, M., Heinze, H.-J., and Rotte, M. (2009). My third arm: Shifts in topography of the somatosensory homunculus predict feeling of an artificial supernumerary arm. *Human Brain Mapping*, 30(5):1413–1420.

Shin, M., Lee, S., Song, S. W., and Chung, D. (2021). Enhancement of perceived body ownership in virtual reality-based teleoperation may backfire in the execution of high-risk tasks. *Computers in Human Behavior*, 115:106605.

Škola, F. and Liarokapis, F. (2016). Examining the effect of body ownership in immersive virtual and augmented reality environments. *The Visual Computer*, 32(6):761–770.

Slater, M., Pérez-Marcos, D., Ehrsson, H. H., and Sanchez-Vives, M. V. (2009). Inducing illusory ownership of a virtual body. *Frontiers in Neuroscience*, 3(2):29.

Slater, M., Spanlang, B., Sanchez-Vives, M. V., and Blanke, O. (2010). First person experience of body transfer in virtual reality. *PloS one*, 5(5):e10564.

Snow, M. P. (1998). *Charting presence in virtual environments and its effects on performance*. PhD thesis, Virginia Tech.

Song, H., Kim, J., and Lee, K. M. (2014). Virtual vs. real body in exergames: Reducing social physique anxiety in exercise experiences. *Computers in Human Behavior*, 36:282–285.

Steptoe, W., Julier, S., and Steed, A. (2014). Presence and discernability in conventional and non-photorealistic immersive augmented reality. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 213–218. IEEE.

Steptoe, W., Steed, A., and Slater, M. (2013). Human tails: ownership and control of extended humanoid avatars. *IEEE transactions on visualization and computer graphics*, 19(4):583–590.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4):73–93.

Stevenson Won, A., Bailenson, J., Lee, J., and Lanier, J. (2015). Homuncular flexibility in virtual reality. *Journal of Computer-Mediated Communication*, 20(3):241–259.

Stone, K. D., Bullock, F., Keizer, A., and Dijkerman, H. C. (2018). The disappearing limb trick and the role of sensory suggestibility in illusion experience. *Neuropsychologia*, 117:418–427.

Suzuki, K., Garfinkel, S. N., Critchley, H. D., and Seth, A. K. (2013). Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. *Neuropsychologia*, 51(13):2909–2917.

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

Toet, A., Kuling, I. A., Krom, B. N., and Van Erp, J. B. (2020). Toward enhanced teleoperation through embodiment. *Frontiers in Robotics and AI*, 7(14).

Tsakiris, M. (2010). My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia*, 48(3):703–712.

Tsakiris, M., Carpenter, L., James, D., and Fotopoulou, A. (2010a). Hands only illusion: Multisensory integration elicits sense of ownership for body parts but not for non-corporeal objects. *Experimental Brain Research*, 204(3):343–352.

Tsakiris, M. and Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1):80–91.

Tsakiris, M., Longo, M. R., and Haggard, P. (2010b). Having a body versus moving your body: neural signatures of agency and body-ownership. *Neuropsychologia*, 48(9):2740–2749.

Tsakiris, M., Prabhu, G., and Haggard, P. (2006). Having a body versus moving your body: How agency structures body-ownership. *Consciousness and cognition*, 15(2):423–432.

Van Erp, J. B., Van Veen, H. A., Jansen, C., and Dobbins, T. (2005). Waypoint navigation with a vibrotactile waist belt. *ACM Transactions on Applied Perception*, 2(2):106–117.

Vella, D. (2013). "It's a-me/Mario": Playing as a ludic character. In *Foundations of Digital Games Conference Proceedings*, pages 31–38.

Venables, P. and Christie, M. (1980). *Techniques in Psychophysiology*, chapter Electrodermal activity, pages 2–67. Addison-Wesley.

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

Weibel, D. and Wissmath, B. (2011). Immersion in computer games: The role of spatial presence and flow. *International Journal of Computer Games Technology*, 2011:6.

Weibel, D., Wissmath, B., and Mast, F. W. (2010). Immersion in mediated environments: the role of personality traits. *Cyberpsychology, Behavior, and Social Networking*, 13(3):251–256.

Welch, R. B. (1999). How can we determine if the sense of presence affects task performance? *Presence*, 8(5):574–577.

Wen, W., Yamashita, A., and Asama, H. (2015). The sense of agency during continuous action: performance is more important than action-feedback association. *PloS one*, 10(4):e0125226.

Wilhelmsson, U. (2001). *Enacting the point of being-computer games, interaction and film theory: affordances and constraints, metaphorical concepts and experientialist cognition observed through the environment in computer games.* PhD thesis, University of Copenhagen, Denmark.

Wilhelmsson, U. (2006). What is a Game Ego: or How the embodied mind plays a role in computer game environments. *Affective and emotional aspects of human-computer interaction: Game-based and innovative learning approaches*, 1:45–58.

Wilhelmsson, U. (2008). Game Ego presence in video and computer games. In *Extending Experiences. Structure, analysis and design of computer game player experience*, pages 58–72. Lapland University Press, Rovaniemi, Finland.

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., Saari, T., Laarni, J., Ravaja, N., Gouveia, F. R., et al. (2007). A process model of the formation of spatial presence experiences. *Media Psychology*, 9(3):493–525.

Witmer, B. G. and Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240.

English Summary

Augmented reality (AR) is a new and upcoming medium that promises innovations in many application areas, ranging from entertainment games to serious teleoperations. Whereas virtual reality (VR) uses completely virtual environments, AR uses a combination of real and virtual to create one seamlessly mixed environment. This makes AR a technologically, perceptually, and cognitively complex area, which coincides with a lack of research on theoretical and experiential aspects of AR. With an eye towards the growing AR gaming market, this dissertation makes first steps in those less explored directions, namely in the area of multimodal game experiences that rely on the body for interaction. Specifically, I examine how supernumerary virtual avatars are experienced when presented alongside the player's real body.

This research had two main starting points. The first was to define multimodal AR, since nearly all previous AR research has solely focused on visual AR, and the corresponding definitions could not easily be translated to multimodality. I defined the components of multimodal AR to disambiguate the terms 'real' and 'virtual' and created a classification system for multimodal AR based on the combination of relevant modalities. The second starting point was to create an overview of player-avatar links in games from an interdisciplinary perspective. The presented works from Humanities, Natural and Social Sciences, showed a clear split in views: those that argue that the link must be related to embodiment, and those who argue it is more than embodiment. Focusing on embodiment views specifically, it became clear that the common reasoning did not encompass all forms of avatar presentation possible in AR, for example when the own player's body was present alongside (part of) the avatar's body. As a result, it was not clear whether embodiment of these avatars was even feasible.

Based on this foundation, a series of experiments was conducted to study the feasibility and function of this supernumerary embodiment in AR. In particular,

I found that there were certain requirements for the experience of embodiment with respect to multimodality and realism that were related to the dissonance between real and virtual in AR, and that this experience may have been influenced by participants' capability to become immersed in media in general. Taking these findings into account, I presented a two-level processing model for this specific form of media-related embodiment that suggests that such media-related variables are involved in a processing level after top-down and bottom-up signal processing. In a final experiment, I studied the practical benefits of embodiment in AR by investigating the relation between embodiment and task performance, in a task involving a third virtual arm. The results showed that there was an indirect positive relation between the feeling of owning the third arm and performance that was mediated by the feeling of agency over the third arm.

Nederlandse Samenvatting

Augmented reality (AR) is een nieuw en aanstaand medium dat innovaties belooft in veel toepassingsgebieden, van entertainment games tot serieuze teleoperaties. In tegenstelling tot virtual reality die een compleet virtueel omgeving gebruikt, gebruikt AR een combinatie van echt en virtueel om een naadloos gemengde omgeving te realiseren. Dit maakt AR een technologisch, perceptueel en cognitief complex onderzoeksgebied, wat samenvalt met een gebrek aan onderzoek op het gebied van theoretische en ervaringsaspecten van AR. Met het oog op de groeiende AR gaming markt maakt dit proefschrift de eerste stappen in deze minder verkende richtingen, namelijk op het gebied van multimodale game ervaringen die gebruik maken van het lichaam voor interactie. Ik onderzoek met name hoe boventallige virtuele avatars ervaren worden wanneer deze tegelijk worden gepresenteerd met de spelers echte lichaam.

Dit onderzoek had twee uitgangspunten. Het eerste was om multimodale AR te definiëren, gezien bijna alle eerdere AR onderzoek alleen op visuele AR toegespitst was, en bijbehorende definities niet eenvoudig vertaald kunnen worden naar multimodaliteit. Ik definieer de componenten van multimodale AR om de termen ‘echt’ en ‘virtueel’ ondubbelzinnig te maken en vormde een classificatiesysteem voor multimodale AR gebaseerd op de combinatie van relevante modaliteiten. Het tweede uitgangspunt was om een overzicht van speler-avatar relaties in games te maken vanuit een interdisciplinair perspectief. De gepresenteerde artikelen uit de Geestes-, Beta- en Sociale Wetenschappen toonden een duidelijke verdeling van standpunten: zij die beweren dat de link wel gerelateerd moet zijn aan belichaming, en zij die beweren dat het meer moet zijn dan belichaming. Door te focussen op alleen belichamingsstandpunten, werd het duidelijk dat de algemene redenering niet alle vormen van avatar presentatie omvatte die mogelijk zijn in

AR, bijvoorbeeld wanneer de spelers eigen lichaam tegelijk aanwezig was met (een deel van) het lichaam van de avatar.

Gebaseerd op het voorgaande, is een serie experimenten uitgevoerd om de haalbaarheid en functionaliteit van deze boventallige belichaming in AR te bestuderen. In het bijzonder bleken er bepaalde vereisten te zijn wat betreft de dissonantie tussen echt en virtueel in AR, en dit had mogelijk te maken met in hoeverre de deelnemers in staat waren tot zich in te leven in media in het algemeen. Inachtname van deze bevindingen, presenteerde ik een 2-stap verwerkingsmodel voor deze specifieke vorm van media gerelateerde belichaming dat suggereert dat zulke media gerelateerde variabelen betrokken zijn in een verwerkingsstap na top-down en bottom-up verwerking. In een laatste experiment bestudeerde ik de praktische voordelen van belichaming in AR door de relatie tussen belichaming en taak prestatie te onderzoeken, in een taak met een derde virtuele hand. De resultaten toonden aan dat er een indirecte positieve relatie was tussen het eigendomsgevoel over de derde hand en prestatie, en dat deze bemiddeld werd door het controlegevoel over die hand.

Curriculum Vitae

Nina Rosa holds two bachelor degrees in Mathematics and Computer Science, and a cum laude master degree in Game and Media Technology, all from Utrecht University. In her interdisciplinary PhD research at Utrecht University, she studies supernumerary embodiment in multimodal augmented reality games. She was a member of the Natural Science Faculty Council (2017-2018) at Utrecht University.

Acknowledgements

I want to thank my supervisors Wolfgang Hürst, Remco Veltkamp and Peter Werkhoven. From the start you all provided great guidance in this project, and still gave me so much freedom to study those areas that intrigued me.

Wolfgang, thank you for your frequent and extremely detailed feedback.

Remco, thank you for helping me get as much out of this PhD as I could.

Peter, thank you for always making time for me in your crazy schedule.

Thanks to my kids, for reminding me about the important things in life. Did you help delay the writing of this dissertation by a few years? Possibly. But you also helped me want to create a better future, and helped me find my passion for sustainability.

I also want to thank my husband, Jelmer, for keeping me sane during these past six years. You would not hesitate to help me test every part of an experiment in the days before my first participant, or run the house while I was working on sleepless paper deadlines. You accompanied me to conferences abroad, and to the job interview that got me this very PhD position.

Finally, thanks to all the students, scientists, engineers, family and friends that contributed to this research.