Designing for Hand Ownership in Interaction with Virtual and Augmented Reality

Tiare Feuchtner

PhD Dissertation

Department of Computer Science
Aarhus University
Denmark
Designing for Hand Ownership in Interaction with Virtual and Augmented Reality

A Dissertation Presented to the Faculty of Science and Technology of Aarhus University in Partial Fulfillment of the Requirements for the PhD Degree

by
Tiare Feuchtner
June 30, 2018
Abstract

This thesis explores the body ownership illusion during interaction in virtual and augmented reality; in particular, how the body ownership illusion can be leveraged to overcome the limitations of our physical bodies and to create more efficient and engaging interactions. Body ownership is the feeling that we inhabit a physical body through which we experience and interact with the world. The conviction that your physical hand is, in fact, your own hand is an example of body ownership. Under certain conditions, this feeling can also be elicited of an artificial body or body part (e.g., a rubber hand or virtual hand), which is then known as the body ownership illusion.

The research questions presented in this thesis were explored with two independent prototypes: the first consisting of an augmented reality system and the second providing a virtual reality experience. The state-of-the-art for immersive virtual reality involves interaction with the virtual world through controllers that we perceive as tools hovering in mid-air. We cannot see our hands gripping these tools, nor do we have any other indicators of our body in our peripheral vision or through the presence of shadows. This ghostly, body-less state is unnatural, and a constant reminder that the world we see around us is not real. Our sense of presence is limited if we do not have a body that inhabits part of the environment we perceive. In augmented reality applications on the other hand, we typically perceive our own body as part of the real world and can sometimes use our physical hands to manipulate virtual content. In contrast to virtual hands, however, our physical hands cannot be flexibly adapted for a specific task (e.g., extending our reach). Hence, I explore how we can create a virtual body representation to which the user feels strongly connected through the body ownership illusion. I argue that this inherently changes the quality of the interaction; while traditionally the screen forms a clear boundary between the real and the virtual world, the illusion of owning a virtual body moves this border into our own minds. In other words, a virtual body representation may seem just as real to us as our physical bodies, so we become part of the virtual world.

The findings presented in this thesis show that we can control an unnaturally long virtual arm that we perceive as part of our body. Furthermore, we can feel body ownership of a virtual hand that appears to be reaching towards an overhead target, while our physical hand rests comfortably at waist-level. We may even feel body ownership of a very unrealistic, disconnected virtual hand, if it follows our movements naturally. The results thus indicate that our mental body representation may be more malleable than was previously believed. Consequently, while most approaches to interaction design attempt to modify the virtual world or the tools with which we interact, we can instead augment the body we interact through - or at least our mental image thereof.
Resumé


Resultaterne præsenterer i denne afhandling viser, at vi kan kontrollere en unaturligt lang virtuel arm, som vi oplever som en del af vores kroppe. Ydermere, kan vi opleve body ownership over en virtuel hånd, der rækker efter et mål over vores hoveder, alt imens vores fysiske hænder hviler komfortabelt i taljehøjde. Vi kan endda opleve body ownership over en meget urealistisk og afkoblet virtuel hånd, såfremt den følger vores bevægelser naturligt. Resultaterne viser således, at vores mentale kropsforståelse er mere formbar og påvirkelig end hidtil antaget. Mens de fleste tilgange til interaktionsdesign
forsøger at ændre den virtuelle verden, eller værktøjene vi interagerer med, kan vi i stedet ændre eller manipulere den krop, vi interagerer igennem - eller i det mindste vores mentale billede deraf.
Acknowledgments

Thanks are due first and foremost to my supervisors Jörg and Kaj for giving me this opportunity and the support I needed to reach my goal. Jörg, you saw the researcher in me long before I did - thank you for being a great mentor and friend throughout all these years.

I would also like to thank my colleagues here at AU who offered not only their professional feedback and help, but also their friendship. I wish to thank Andreas F. for joining me on this adventure, which kind of already started in Berlin. Special thanks to Allan, Andreas F., Andreas M., Anke, Emil, Eve, Henrik, Ida, Jörg, Kasper, Lindsay, Mads, Majken, Matthias, Matus, Mille, Mirzel, Myroslav, Natalie, Peter, Sarah, Tim, Tobias, and Troels for participating in my user studies, and to Eve, Matthias, and Tobias for help with practical things like figuring out electronics.

To my colleagues at the EventLab I am grateful for welcoming me into their group and making my stay abroad so memorable. Special thanks to Mel for making it possible in the first place, and to Anna, Birgit, Irene, Marta, Solene, and Xavi for help with organizational and technical aspects of my experiment. I also wish to thank Alejandro, Anna, Birgit, Dani, Domna, Justyna, Marta, Mel, Ramon and Sameer for their participation in my pilot studies and for excellent feedback.

I am deeply grateful to Emily, Michael and Andreas F. for editing this thesis and to Ida, Matthias and Mirzel for help with the Danish abstract.

Furthermore, my expat-family must not be forgotten. Andrea & Tibor, Emily & Josh, and Laura & Leendert, you made Denmark feel like home to me. Thanks to all the great friends (there are too many to be named) who joined me for barbecues in the rain, bike rides, city trips, days at the beach, and game nights. Skål to all the fun times we had - may there be many more!

I also want to thank my family: inexplicably, you have always been proud of me for doing things you don’t quite understand. I am specially grateful to my mother and my brother Anjin, for always supporting me without questioning my decisions, and for not letting me lose sight of what is truly important.

Last but not least, I can’t even begin to say how much I owe to Michael. Thank you for your tireless support and for being my most patient and enduring pilot participant. Thank you for listening to all my great and also the not-so-great ideas, for challenging me, and for providing an endless supply of much-needed hugs. I could not have done this without you.

Tiare Feuchtner,
Aarhus, June 30, 2018
Structure of the dissertation

This thesis consists of two parts. Part I presents two systems that were designed and developed in the course of my PhD, as well as the research methods applied in the experiments. These experiments are reported on in three main papers which are included in Part II of this dissertation. These papers are mainly referred to by their short names (Paper A: LongArm, Paper B: Ownershift, Paper C: PhysioMotor).

Part I is structured as follows:

Chapter 1 (Introduction) describes the motivation, presents the research questions and provides the context for this research. Furthermore, it outlines the structure of Part I.

Chapter 2 (Background) presents related work on interaction techniques, fatigue and performance, as well as topics like cognition and self-perception, embodiment and the body ownership illusion. It also discusses the limitations of, and requirements for, eliciting the body ownership illusion.

Chapter 3 (System design and implementation) provides general details about the design process and implementation of the systems that were developed to support the presented research.

Chapter 4 (Augmented reality system for body ownership) presents the augmented reality (AR) system in detail. This system served as platform to explore the virtual hand illusion (VHI) in AR. By describing specific implementation aspects that were found critical for the body ownership illusion, this section may serve as guide for future researchers who wish to elicit the feeling of body ownership of a virtual human body in AR. Detailed results of the user study are presented in the LongArm [34] publication, attached in chapter 8.

Chapter 5 (Virtual reality system for body ownership) presents details about the virtual reality (VR) system, with which the experiments presented in Ownershift [35] and PhysioMotor [36] were conducted. This chapter includes a description of all components needed for recreating the VHI in immersive VR, as well as considerations for ownership-preserving interaction design. Empirical study results can again be found in the respective publications, in chapters 9 (Ownershift) and 10 (PhysioMotor).

Chapter 6 (Research approach) gives details on the experimental procedure for user studies, study design, and methods for data collection.
Information about the data analysis and results may be found in the respective publications attached in Part II.

Chapter 7 (Conclusion) offers a summary of the main results, revisits the research questions, and highlights the contributions of the presented work. This is followed by a discussion of challenges and directions for future research.

The papers included in Part II of this thesis are shortly described by their titles and abstracts as follows.


In this paper, we explore how users can control remote devices with a virtual long arm, while preserving the perception that the artificial arm is actually part of their own body. Instead of using pointing, speech, or a remote control, the users' arm is extended in augmented reality, allowing access to devices that are out of reach. Thus, we allow users to directly manipulate real-world objects from a distance using their bare hands. A core difficulty we focus on is how to maintain ownership for the unnaturally long virtual arm, which is the strong feeling that one’s limbs are actually part of the own body. Fortunately, what the human brain experiences as being part of the own body is very malleable and we find that during interaction the user’s virtual arm can be stretched to more than twice its real length, without breaking the user’s sense of ownership for the virtual limb.


We present Ownershift, an interaction technique for easing overhead manipulation in virtual reality, while preserving the illusion that the virtual hand is the user’s own hand. In contrast to previous approaches, this technique does not alter the mapping of the virtual hand position for initial reaching movements towards the target. Instead, the virtual hand space is only shifted gradually if interaction with the overhead target requires an extended amount of time. While users perceive their virtual hand as operating overhead, their physical hand moves gradually to a less
strained position at waist level. We evaluated the technique in a user study and show that Ownershift significantly reduces the physical strain of overhead interactions, while only slightly reducing task performance and the sense of body ownership of the virtual hand.


We elicit the virtual hand illusion through different multisensory stimuli and explore how these modulate body ownership of realistic and unrealistic hands. We compare visual-motor, visual-proprioceptive and cardio-visual stimuli to elicit body ownership of three different virtual hand representations: a very realistic hand, an abstract hand model, and a very unrealistic, disconnected hand representation. For visual-motor synchrony we support full hand tracking, allowing natural movement of the virtual hand. The visual-proprioceptive and cardio-visual conditions are so-called vision-only conditions (no hand tracking). In the latter the participant’s heartbeat is visualized on the virtual hand. Our results show that realism of the virtual hand significantly affects the body ownership illusion. We found no significant differences between the vision-only and motor conditions, nor was there an influence of the heartbeat visualization. However, 50% of participants indicated body ownership even of very unrealistic hands, if they could move the hand.
## List of acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D</td>
<td>three-dimensional</td>
</tr>
<tr>
<td>AR</td>
<td>augmented reality</td>
</tr>
<tr>
<td>BAQ</td>
<td>body awareness questionnaire</td>
</tr>
<tr>
<td>BOI</td>
<td>body ownership illusion</td>
</tr>
<tr>
<td>ECG</td>
<td>electrocardiogram</td>
</tr>
<tr>
<td>HCI</td>
<td>human computer interaction</td>
</tr>
<tr>
<td>HMD</td>
<td>head mounted display</td>
</tr>
<tr>
<td>IS</td>
<td>interoceptive sensitivity</td>
</tr>
<tr>
<td>LRA</td>
<td>linear resonant actuator</td>
</tr>
<tr>
<td>PC</td>
<td>personal computer</td>
</tr>
<tr>
<td>RHI</td>
<td>rubber hand illusion</td>
</tr>
<tr>
<td>RMSE</td>
<td>root-mean-square error</td>
</tr>
<tr>
<td>VHI</td>
<td>virtual hand illusion</td>
</tr>
<tr>
<td>VHS</td>
<td>virtual hand space</td>
</tr>
<tr>
<td>VR</td>
<td>virtual reality</td>
</tr>
</tbody>
</table>
Contents

Abstract i

Resumé iii

Acknowledgments v

Structure of the dissertation vii

List of acronyms xi

Contents xiii

I Overview 1

1 Introduction 3
   1.1 Motivation and research questions 4
   1.2 Context of research: body ownership 6
   1.3 Platforms for experimentation 7
   1.4 Structure of Part I 9

2 Background 11
   2.1 Cognition and self-perception 11
   2.2 Body ownership illusions 14
   2.3 From reality to virtuality 19
   2.4 Concepts for interaction in virtual reality 21
   2.5 Positioning of own work 26

3 System design and implementation 29
   3.1 Iterative design and development 30
   3.2 Designing for body ownership 31

4 AR system for body ownership 33
   4.1 AR system components 33
   4.2 Elicitation of body ownership in AR 37
   4.3 Implementation details 39

5 VR system for body ownership 47
   5.1 VR system components 48
   5.2 Elicitation of body ownership in VR 54
   5.3 Implementation details 57

6 Research approach 65
   6.1 Experimental procedure for user studies 65
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2 Study design</td>
<td>67</td>
</tr>
<tr>
<td>6.3 Data collection</td>
<td>71</td>
</tr>
<tr>
<td>7 Conclusion</td>
<td>81</td>
</tr>
<tr>
<td>7.1 Conclusions of LongArm (chapter 8)</td>
<td>81</td>
</tr>
<tr>
<td>7.2 Conclusions of Ownershift (chapter 9)</td>
<td>81</td>
</tr>
<tr>
<td>7.3 Conclusions of PhysioMotor (chapter 10)</td>
<td>82</td>
</tr>
<tr>
<td>7.4 Revisiting the research questions</td>
<td>82</td>
</tr>
<tr>
<td>7.5 Limitations</td>
<td>84</td>
</tr>
<tr>
<td>7.6 Future work</td>
<td>84</td>
</tr>
<tr>
<td>7.7 Outlook</td>
<td>85</td>
</tr>
</tbody>
</table>

## II Publications

### 8 LongArm publication

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 Introduction</td>
<td>90</td>
</tr>
<tr>
<td>8.2 Related Work</td>
<td>91</td>
</tr>
<tr>
<td>8.3 The Psychological Concept of Ownership</td>
<td>93</td>
</tr>
<tr>
<td>8.4 Supporting Ownership in Interaction</td>
<td>95</td>
</tr>
<tr>
<td>8.5 Design Process</td>
<td>99</td>
</tr>
<tr>
<td>8.6 Evaluation</td>
<td>100</td>
</tr>
<tr>
<td>8.7 Results</td>
<td>104</td>
</tr>
<tr>
<td>8.8 Discussion</td>
<td>108</td>
</tr>
<tr>
<td>8.9 Conclusion</td>
<td>110</td>
</tr>
</tbody>
</table>

### 9 Ownershift publication

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1 Introduction</td>
<td>112</td>
</tr>
<tr>
<td>9.2 Related Work</td>
<td>113</td>
</tr>
<tr>
<td>9.3 Ownershift</td>
<td>117</td>
</tr>
<tr>
<td>9.4 System design</td>
<td>118</td>
</tr>
<tr>
<td>9.5 Experiment</td>
<td>119</td>
</tr>
<tr>
<td>9.6 Results</td>
<td>124</td>
</tr>
<tr>
<td>9.7 Discussion</td>
<td>131</td>
</tr>
<tr>
<td>9.8 Conclusion</td>
<td>132</td>
</tr>
<tr>
<td>9.9 Appendix</td>
<td>133</td>
</tr>
</tbody>
</table>

### 10 PhysioMotor publication

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.1 Introduction</td>
<td>135</td>
</tr>
<tr>
<td>10.2 Background</td>
<td>137</td>
</tr>
<tr>
<td>10.3 Experimental design</td>
<td>138</td>
</tr>
<tr>
<td>10.4 Materials and methods</td>
<td>141</td>
</tr>
<tr>
<td>10.5 Conclusion</td>
<td>158</td>
</tr>
</tbody>
</table>

**Bibliography** | 159 |
Part I

OVERVIEW
Introduction

In our daily lives most of us use a mouse to control the cursor to interact with a computer system. However, recent advances in technology allow the speculation that immersive virtual reality (VR) and augmented reality (AR) systems are the computer systems of the near future. For direct manipulation in immersive VR environments the current state-of-the-art still involves a kind of cursor; namely a virtual representation of the controller being used (e.g., the Vive\textsuperscript{1} controller depicted in Fig 1.1, A, B). The visual representation of this cursor may be adapted to better fit its purpose in different environments, such as representing a gun and shield in the Space Pirate Trainer\textsuperscript{2} (Fig 1.1, C). AR applications often involve a hand-held device through which the user perceives and interacts with the virtual augmentations of the real world (e.g., Vuforia Chalk\textsuperscript{3}, Fig 1.1, D). More direct interaction with the virtual content is provided with head-mounted AR systems, such as the Microsoft Hololens\textsuperscript{4} (Fig 1.1, E), where user’s can interact with virtual content directly using their hands. However, interactions in these systems are frequently limited to a small set of command gestures. The recently announced Project North Star\textsuperscript{5} by LeapMotion takes a step towards more direct manipulation of the virtual world, as shown in Fig 1.1 (F).

While visually adapted cursors are excellent for communicating their functionality, my view is that such cursors should only be presented when the use of a hand-held tool is justified by the task (e.g., using a brush to paint with, or holding a gun to shoot). In most other cases, the users should be allowed to interact with their hands alone, i.e., without the use of a controller. In both categories, when using a virtual tool or interacting with bare hands, the users should be provided with a virtual body representation that is connected to the virtual hand through which they interact. In contrast to physical hands, this virtual hand can then be flexibly adapted and transformed to suit the application (e.g., stretching the virtual arm to extend our reach).

The objective of my work is to create a strong connection between the users and their virtual body representation. I aim to create this connection through the body ownership illusion, which not only changes what we think we can control, but even what we believe to be. In other words, it is my goal to let users interact with the virtual (or virtually augmented) world through a virtual arm, which they perceive as part of their physical body.

\textsuperscript{1}Vive: \url{https://www.vive.com} (last accessed: 28.06.2018)
\textsuperscript{2}Space Pirate Trainer: \url{http://www.spacepiratetrainer.com} (last accessed: 28.06.2018)
\textsuperscript{3}Vuforia Chalk: \url{https://chalk.vuforia.com} (last accessed: 28.06.2018)
\textsuperscript{5}Project North Star by LeapMotion: \url{https://developer.leapmotion.com/northstar} (last accessed: 28.06.2018)
Chapter 1. Introduction

Figure 1.1: State-of-the-art for interaction in VR and AR. When interacting with VR using Vive controllers (A), these are also displayed in the virtual environment (B). The visual representation of these controllers can further be adjusted to suit the application, e.g., as a gun and shield in a space shooter (C). AR solutions often involve a hand-held device with touch-screen (D), or head-mounted displays. Here virtual content can be manipulated with the hands through predefined gestures (E), or more direct manipulations (F).

1.1 Motivation and research questions

One of the main differences between traditional interactions with computers (through mouse and keyboard) and interaction in immersive VR is the presence of our own body. While we perceive our own physical hand moving the mouse and can clearly observe the mapping between the movements of the mouse and the cursor on the screen, the state-of-the-art for VR involves a head mounted display (HMD) through which we peer into the virtual world and interact in a space where all reference to our body is removed. We know of course that the seemingly floating cursors (such as paintbrushes or guns) represent the controllers that we are holding in our hands, and we feel in control in this virtual world. Disembodied users may feel very present in the

virtual environment, but once they experience a virtual body that faithfully follows their movements, they must realize that there are different levels of presence [57, p.47]. I argue that if there were a virtual body representation, which we perceived to be our own (body ownership), this would make the experience feel much more ‘real’. It would transform us from being observers and actors-by-proxy (i.e., through controllers) to being part of the virtual world itself.

Apart from the experiential aspect of interaction, using our own (virtual) hands to interact with the system may also minimize the need for training and reduce cognitive load. I further identify three additional points in which perceived body ownership of a virtual user representation, and in particular virtual hands, can be beneficial:

1. Locating the end-effector: If we want to point at something with a mouse cursor, we first have to locate the cursor on the screen. This occasionally involves frantic shaking of the mouse to make the cursor noticeable. In contrast, when reaching for something with our hands we can begin the reaching movement right away, since our proprioception constantly tells us where our hands are. Similarly there is no need for locating our (own) virtual hand representations before initiating interaction. I argue that this is even the case when the virtual hand space is transformed, leading to a discrepancy between the position of our physical hand and the virtual one, since our proprioception is updated to incorporate the artificial limb that we perceive as part of our body.

2. Immediate usability: we already know how to use our hands, so there is no need to learn new controls. For example, if we want to grab something and flip it around, we don’t have to learn to use the correct tool, manipulation point, or key combination to do this. If properly supported by the computer system, we could simply reach for a virtual object with our hand and manipulate it just as we have learned to manipulate physical objects in everyday life.

3. Preventing errors: we are aware of the limitations of what we can do with our hands and the dangers to them, so there is no need for creating new rules or restrictions. For example, we know we can’t reach through solid walls and that touching a hot surface, or a sharp edge, will hurt. Applying such metaphors in VR could be used to guide our interactions and prevent mistakes.

Furthermore, interacting with a virtual hand representation potentially allows us to augment our bodies and overcome our physical limitations. For example we could extend our reach (LongArm), or reposition the virtual hand for a
more comfortable body pose during prolonged overhead interaction (Owner-shift). In conjunction with robots (i.e., when the movements of the virtual hand are mapped to a robot arm in the real world) we could develop superhuman strength, manipulate dangerous materials, work remotely in lethal environments, and shrink or grow our workspace to manipulate microscopic elements or large scale construction components, while always perceiving that we are doing this with our own two hands. My vision is to create non-invasive virtual augmentations of our bodies that do not disrupt the integrity of our internal body representation. In pursuing this vision, the following research questions were to be answered:

RQ1 Can the illusion of owning of a virtual hand be elicited in AR as well as VR? (LongArm)

RQ2 Can body ownership of a virtual hand be maintained, if this hand is being actively used to manipulate objects? In particular, if unnatural transformations are applied to the virtual hand representation, such as:

- the user can actively stretch the virtual arm to supernatural length to reach distant objects (LongArm)
- the virtual hand is shifted upward in relation to the user’s real hand, allowing overhead interaction in a comfortable posture (Owner-shift)

RQ3 Can users manipulate their environment with a transformed virtual hand (e.g., unnatural length or position) without extensive training? (LongArm, Owner-shift)

RQ4 To what degree does the type of multisensory stimuli (visuo-motor, visuo-propiroceptive, cardio-visual) impact the body ownership illusion, particularly for unrealistic hand representations? (PhysioMotor)

1.2 Context of research: body ownership

Our bodies are the tangible, visible representations of our ‘self’ in the physical world. We perceive our bodies as our ‘own’ and this feeling is referred to as body ownership [125]. Any healthy person would have no doubt that the hand attached to their right arm is in fact their own right hand and not someone else’s, or part of their environment.

However, research has shown that our self-perception is malleable, and what we perceive as our body can be changed [125], for instance through the body ownership illusion. The body ownership illusion refers to the perception that an artificial body, or part thereof, is (part of) the own physical body. This manifests itself through the expectation of being able to move the body (part)
1.3. PLATFORMS FOR EXPERIMENTATION

as you would move your own, and by experiencing external stimuli with the surrogate body (part) such as temperature, touch, and pain.

The key to evoking the illusion of body ownership of a virtual body representation is multisensory integration [20, 80]. Multisensory integration refers to the combination of stimuli from multiple senses to form a consistent perception of the world. For instance, we look around to visually locate a sound source, or we might touch a metallic-looking surface to discern whether it is metal or plastic. It is through the integration of such stimuli that we are also prone to illusions, such as the ventriloquist effect [4], or the above mentioned body ownership illusion.

The most commonly known body ownership illusion is the Rubber Hand Illusion (RHI) [15], where a person is made to perceive a rubber hand as part of the physical body. This is achieved by applying tactile stimuli to the rubber hand, which is being observed by the person, and at the same time stimulating the person’s real hand, which is hidden from view, in exactly the same way (e.g., stroking the fingers with a brush). This is illustrated in Fig 1.2 (A). Since the stimuli a person sees delivered to the artificial hand match exactly with the touch felt on the physical hand, the person then comes to perceive the rubber hand as part of the own body (Fig 1.2, B).

Initially the body ownership illusion was studied mainly within cognitive psychology with the purpose of better understanding how humans perceive the world, how sensory input is processed in the brain, and how we form and update our internal body representation. Soon it became apparent that the illusion of owning an artificial body, or part thereof, could be beneficial for therapy of phantom limb pain, where a similar principle was already applied in mirror therapy. Further potential medical uses include rehabilitation, physiotherapy, and the treatment of anxiety [24, 44, 93].

1.3 Platforms for experimentation

The three publications discussed in this thesis explore boundaries of the body ownership illusion and investigate approaches that leverage this illusion to aid interactions with both our virtual and real environments. VR is a powerful tool for exploring the illusion of body ownership since it provides both a controlled environment, and also a high degree of control over the user’s body representation.

Two distinct prototypes served as platforms for experimentation. The first is an augmented reality system (chapter 4), which supports direct manipulation of distant real-world objects with a long virtual arm. To my knowledge, the publication based on this system (LongArm [34], chapter 8) is the first work to explore the body ownership illusion in AR (if we neglect studies with projector-based systems [54]). The second prototype provides an immersive
CHAPTER 1. INTRODUCTION

Figure 1.2: The Rubber Hand Illusion (RHI). In the RHI experiment the participant’s hidden real hand and a rubber hand are simultaneously stroked with a brush (A). Feeling touch that coincides with the strokes observed on the rubber hand leads to the illusion that touch can be felt on the rubber hand. This in turn suggests the explanation that the rubber hand is part of the own body. It thereby replaces the real hand in our internal body representation (B).

virtual reality application (chapter 5) in which users interact with virtual targets using a virtual hand representation. This system enabled the design and evaluation of an interaction technique that allows manipulation of overhead targets in a comfortable pose by slowly moving the virtual hand towards the targets (Ownershift [35], chapter 9). Furthermore, it served to explore the effect of different combinations of multisensory stimuli (visuo-proprioceptive, visuo-motor, cardio-visual) and different virtual hand representations (realistic/unrealistic) on the body ownership illusion (PhysioMotor [36], chapter 10).

The computer systems discussed immerse the users in a virtual environment or a virtually augmented version of their real environment. The users experience this environment through a head mounted display and can interact with it using nothing but their own hands. A network of sensors and actuators, either in the environment or wearable on the body, keep track of the users to enable interaction and provide feedback.
1.4 Structure of Part I

Part I of this thesis is structured as follows: the next chapter (chapter 2) will provide the background for this thesis, discussing cognition and perception, the body ownership illusion, the virtuality continuum, computer interaction in general and some specific interaction techniques. This is followed by a chapter which gives general information on my approach for system design and implementation (chapter 3). Chapters 4 and 5 give a more detailed description of the AR and VR systems respectively, including design considerations and implementation details that are not contained in the publications. As presented in table 1.1, LongArm features an exploration of the research questions based on the AR system, whereas Ownershift and PhysioMotor are based on the VR platform. Chapter 6 describes the methodology and experimental procedure. The research focus for LongArm and Ownershift was to explore body ownership illusions (BOIs) as a quality of interaction, whereas PhysioMotor presents more fundamental research on BOIs, as is indicated by the columns of table 1.1. Finally, a discussion of the main results, conclusions and directions for future work are given in chapter 7.

<table>
<thead>
<tr>
<th>Platform</th>
<th>Research Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Reality</td>
<td><strong>Paper A</strong> - LongArm: Extending the Body for Interaction with Reality [34]</td>
</tr>
<tr>
<td>(chapter 4)</td>
<td>(chapter 8)</td>
</tr>
<tr>
<td>(chapter 5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Paper C</strong> - PhysioMotor: Exploring different Multi-sensory Stimuli and Visual Realism for the Virtual Hand Illusion [36] (chapter 10)</td>
</tr>
</tbody>
</table>

Table 1.1: **Research focus and exploration platform.** The papers presented in this thesis can be categorized by their research focus and platform that was used for exploration. The focus was either the BOI itself, or the BOI as aspect of interaction. The platforms include AR and VR.

Throughout this thesis the term ‘we’ is used repeatedly when reporting details
CHAPTER 1. INTRODUCTION

about the systems and experiments. In doing so, I refer to the authors of the papers attached in Part II, unless explicitly stated otherwise.
Background

I see my research as interdisciplinary [27], by providing results and insights mainly in the domain of interaction design and additionally exploring concepts and applying methods within a second domain. This second domain is behavioral and cognitive psychology, where the body ownership illusion (BOI) is of particular interest.

This chapter presents research that forms the background for my work, from topics such as cognition and self perception, BOIs, and in particular the virtual hand illusion (VHI). I will also discuss the limitations and requirements for eliciting the BOIs. I furthermore briefly define virtual reality and related terms, and discuss related work on direct manipulation and 3D interaction techniques, distant reaching and fatigue of mid-air interactions.

Some parts of this chapter may be found in very similar form in the publications included in Part II.

2.1 Cognition and self-perception

Cognitive psychology is concerned with how we perceive the world around us and how these sensations are processed. This brings along another interesting question, namely what we perceive as the ‘self’ and how we differentiate between sensations emanating from ourselves and the environment around us. Understanding this is a key part of human computer interaction (HCI), since we must first understand humans in order to design interactions and computer systems for them. Human-centered design is particularly important when designing experiences in virtual and augmented reality [57, p.2].

How we perceive the world

Every day we rely on multisensory integration to match up the multitude of impressions we perceive and make sense of the world around us [20, 80]. For instance, we might look around to locate a sound source in an attempt to make sense of an audible stimulus we are receiving (e.g., a moving car, or a barking dog). Redundancy of sensory input is especially important when the signal from one of our senses is ambiguous. We then attempt to get confirmation through a different channel, or in other words, we integrate the input from multiple senses to create one coherent meaning.

Sometimes this integration of multisensory stimuli leads to misinterpretations, which can lead to a wide range of illusions. An example that is well known from puppet shows is Ventriloquism [4], where sound is perceived to emanate from a different place than its actual source (i.e., from the puppet, instead of
the concealed puppet master). Another such illusion is the McGurk effect\(^1\) [83, 89], which shows that when someone talks the sounds we hear depend on the lip and tongue movements we observe on the speaker. For instance, if we contemplate someone pronouncing the phonemes “ba” and “ga”, we assume that we can easily hear them correctly. However, we do not realize how much we rely on our vision to do this. If we overlay the soundtrack of a person saying “ba” over a video of this person saying “ga”, our brain will wrongly tell us that the phoneme we are hearing is indeed “da”.

Such illusions are not limited to the combination of vision and hearing, but can occur through many different combinations of senses or even individual senses alone. For instance, an unusual arrangement of our limbs can lead us to misjudge the location and number of sources of tactile stimuli. An example is the Aristotle illusion [72] by which we perceive two objects when touching only one with our fingers crossed. This is easily reproduced by crossing your middle and index finger and asking someone else to lightly place a pen between both finger tips, while your eyes are closed. You will most likely be under the impression of feeling touch from two separate pens, since your normal body configuration (having index and middle finger side-by-side) would require two pens to create this type of stimulus. The lack of vision leads to recruitment of previous knowledge about our body and the world to interpret the source of touch. In a similar fashion we can bring forth the so called Pinocchio Illusion, also known as the Phantom Nose Illusion [71, 100]. For this illusion an experimenter instructs a participant (P1) to reach forward and rhythmically tap the nose of a second participant (P2) in front of him. The experimenter taps the nose of P1 in the same rhythm. P1 then begins to feel like he is in fact tapping his own nose, but since his arm is extended while doing so, his nose seems abnormally long (hence the name).

In this manner, multisensory integration is also the key to BOIs, which are further discussed in section 2.2.

**How we perceive our body**

As humans we can recognize our own shadow and mirror image. We also have a good idea of what we look like when we walk even without seeing ourselves, and how much space our body inhabits. Furthermore, the representation of our own body is constantly updated to account for changes [25], which allows us to still recognize ourselves despite wearing a fake mustache and avoid bumping into a door frame while carrying a large object. In literature this body representation is often defined as consisting of two aspects, the body schema and the body image. The body schema allows us to navigate through our environment without bumping into obstacles [80]. It is a kind

---

\(^{1}\)Video by BBC Two “Try The McGurk Effect!”: [https://youtu.be/G-1H8vWm3m0](https://youtu.be/G-1H8vWm3m0) (last accessed: 28.06.2018)
2.1. COGNITION AND SELF-PERCEPTION

of body model, which defines the space we occupy and in which we can act \((peripersonal space)\). This model adjusts to changes, informing us to stoop through the door frame when wearing a tall hat, or what objects we can reach after grabbing a long tool (extended peripersonal space). The body image on the other hand is a kind of mental picture of what our body looks like from the outside \([25]\). This can affect our emotions and behavior. For instance, we can be more communicative when we feel pretty, or more self-confident when we are well dressed.

Perhaps orthogonal to the above, physiological feedback, like the sound of breathing or our heartbeat, is something that we associate with our own body and serves for self-representation \([130]\). It has been shown that perceiving our own physiological feedback (e.g., auditory-vibrotactile feedback of our heartbeat) can also affect our emotions \([121]\), and research has found that self-consciousness stems from the integration of both interoceptive and exteroceptive signals \([7]\). Physiological feedback can nowadays be easily gained through low-cost sensors and unobtrusive wearable devices (e.g., fitness bands and smartwatches), and has therefore found application in diverse areas such as game design \([90]\) and clinical psychology \([102]\).

The sense of embodiment

Apart from perceiving our own body, we are also aware of the effects we have on the world around us. Combined, these aspects are referred to as \(self\)-attribution. A similar term that is gaining popularity, in particular in the context of immersive virtual reality (VR), is \(embodiment\). While this term has multiple meanings, I prefer the definition by which our sense of embodiment depends on three factors: \(body ownership\), \(agency\), and \(self-location\) \([63, 77]\). The intensity of the feeling of embodiment can be evaluated as the combination of its three components.

Body ownership refers to our sense of having one connected body, which we inhabit. For instance the conviction that your right hand is a part of your body with which you feel and interact with the world, means that you to feel body ownership of that hand. Seeing an avatar when looking down at yourself in VR can lead to a sense of body ownership of the virtual body (i.e., a BOI).

Agency refers to the feeling of causing changes in the environment, i.e., when we think ourselves to be the immediate cause of an effect we observe in the world. For instance, we feel agency over the mouse cursor that moves across the screen when we move the mouse accordingly. If the mouse controls were suddenly inverted, our sense of agency would be diminished and we would feel less in control. In immersive VR we also feel a sense of control by being able to look around, navigate through the virtual space, and manipulate virtual objects.
CHAPTER 2. BACKGROUND

The last factor, self-location, refers to where we perceive our “self” to be. Under normal circumstances we perceive our self to reside in our own body, since we look out through our eyes. Immersive VR gives you a first-person perspective of the virtual environment, which leads to a feeling of “being there”, or presence. Our sense of self-location appears to be closely connected to our point of view (e.g., location of our eyes/camera).

There is likely some interplay between these three concepts, however it still remains greatly unclear in what way they are connected. There have been several attempts to disentangle agency and body ownership [17, 136]. Some experiments suggest an effect of body ownership on agency (feeling body ownership makes agency more likely), but not vice-versa, since we can certainly feel agency in absence of body ownership [22, 59, 60] (e.g., we feel responsible for moving an object with a tool, which we do not perceive as part of our body). Furthermore, it seems body ownership depends on self-location, since looking at yourself from a third person perspective has been found to lead to out-of-body illusions instead [61]. Perhaps in summary it could be said that, of the three factors of embodiment, body ownership is the most difficult to achieve.

In the context of VR, the term embodiment is frequently used interchangeably with body ownership. This may be due to the circumstance that a feeling of body ownership often coincides with agency and self-location. However, it is important to understand that we can also experience some degree of embodiment even in absence of body ownership. A more specific term, which is closer to body ownership is self-embodiment: “Self-embodiment is the perception that the user has a body within the virtual world” [57, p.47].

2.2 Body ownership illusions

A body ownership illusion (BOI) refers to the attribution of a body-external object to the own body (also known as ’self-attribution’) [66, 125]. For instance, we can be under the illusion of owning part of the body of a mannequin [15], or a virtual avatar [64, 115], or even an entire artificial body [28, 50, 81, 96, 117]. The most widely known body ownership illusion is the Rubber Hand Illusion [15], described as follows.

The Rubber Hand Illusion

The Rubber Hand Illusion (RHI)\(^2\), first reported by Botvinick and Cohen [15], involves feeling body ownership of a rubber replica of a human hand, which rests on a table next to a subject’s real hand (as shown in Fig 1.2). From

2.2. BODY OWNERSHIP ILLUSIONS

the subject’s perspective the rubber arm is located in a plausible position and orientation, so that it could potentially be part of the own body. The subject’s real hand is hidden from view by a vertical screen, and both the real hand and the rubber hand are stroked with a brush. Since the brushstrokes felt on the own hand match the brushstrokes observed on the rubber hand, this leads to a referral of touch. In other words, the participant is under the impression of feeling the touch delivered to the rubber hand. The logical explanation follows that the rubber hand must therefore be part of the own body - i.e., the participant experiences the illusion of body ownership of the artificial hand.

The Virtual Hand Illusion

The RHI was originally studied using a plastic (rubber) hand that the users perceive as part of their own body. More recently however the effect is frequently explored in a mediated setup, i.e., inducing of the feeling of body ownership of a virtual hand. This is also referred to as the Virtual Hand Illusion (VHI) [6, 54, 64, 73, 105, 115, 116]. The mediated setup brings along several advantages, such as a collocation of the real and virtual hand, which prevents incongruent proprioceptive cues. Furthermore it simplifies the concurrent stimulation of multiple senses. For instance, precise tactile stimuli through vibration motors can be attributed to some visible virtual source (Fig 2.1), and tracking of a user’s hand and finger movements allows realistic animation of the virtual hand in real-time.

Figure 2.1: The virtual hand illusion (VHI). This illusion is the equivalent of the RHI, but in VR and therefore involves body ownership of a virtual hand, or arm. It can for instance be elicited through synchronous visuo-tactile stimulation: Here the experimenter taps the participant’s finger (A) and with each tap the participant sees a virtual ball bounce on the finger of the virtual hand (B).
CHAPTER 2. BACKGROUND

Techniques to elicit body ownership illusions

Extensive exploration of BOIs has lead to many fascinating discoveries about the cognitive processes and the cerebral activities involved [29, 30]. Ehrsson et al. [29] measured activity in the premotor cortex of the human brain, showing that multisensory integration of proprioceptive, visual, and tactile senses, even when caused by an illusion, can lead to bodily self-attribution. At the time this was explained by the theory that vision dominates over proprioception and Longo et al. [76] claimed that RHI is not possible without visuo-tactile stimulation.

However, in a later study Ehrsson et al. [30] evaluated the RHI on blindfolded subjects, showing that the illusion cannot simply be attributed to the dominance of vision over proprioception. In this experiment the subject’s left index finger was moved so that it touched a rubber hand, while at the same time a researcher touched the subject’s right hand, causing the illusion of touching one’s own hand. Furthermore, several experiments confirmed that the illusion could occur without tactile stimulation: Subjects claimed to feel body ownership of a rubber hand by just looking at it for a while [33, 132]. Here the body ownership illusion is likely due to integration of vision and proprioceptive feedback. Similarly, the illusion can occur through visuo-motor correlation (without tactile stimulus), as multiple experiments report BOIs caused by finger or hand movements alone [26, 32, 127, 132]. When subjects can see an artificial hand moving in the same way as they are moving their own hand [26, 69, 105, 127] this is sometimes referred to as the Moving RHI [59]. Kalckert and Ehrsson [59] compared active movement, passive movement, and visuo-tactile stimulation (brush strokes) and reported a similarly strong RHI across all three methods. Jenkinson and Preston [56] studied the moving RHI with a mirror and found that the BOI was enhanced. This suggests that our reflection receives “special treatment” in our minds and can directly be attributed to the self.

It remains debated in how far BOIs depend on multisensory integration alone, or how strongly this bottom-up process is complemented by a top-down process [125]. If we assume bottom-up modulation alone, this would mean that we should be able to embody any object, as long as synchronous multisensory stimuli are given. However, it has been found that the visual and structural characteristics of the surrogate hand (e.g., rubber hand or virtual hand) play an important role. Kalckert & Ehrsson [59] report that the illusion can be caused by any type of multisensory input, given that the artificial body part fulfills some minimal requirements. These limitations are discussed in more detail in section 2.2.

While most research on BOIs focuses on exteroceptive signals (e.g., vision, haptics, proprioception), some authors [7, 120] explored visualization of interoceptive signals, such as the heartbeat. They found that such physiological
feedback, can increase self-identification with and self-location toward a virtual body part [120] or full body [7]. In the latter experiment the participant was recorded with a camera from behind and could see this image of himself through a head mounted display (HMD). The image was then augmented by an outline around the body, which flashed in synchrony with each heartbeat. For comparison, these cardio-visual signals were also applied to a body-sized box, where the effects of increased self-identification and self-location could not be observed. Thus interoceptive signals alone are not sufficient for eliciting the BOI, but can increase the effect if applied to a body-similar object. In this context there appear to be strong inter-personal differences, based on how much subjects are aware of their interoceptive signals, also called their interoceptive sensitivity. Tsakiris et al. [129] discovered that our degree of interoceptive sensitivity is a predictor for the malleability of our internal body representations: people with less interoceptive awareness tend to experience stronger illusions. However, BOIs have also been shown to modulate interoceptive sensitivity, resulting in increased awareness of body-internal processes for participants with low initial interoceptive sensitivity [37].

Flexibility and robustness of body ownership illusions

It has been discovered that BOIs can occur not only for an artificial limb that replaces our own, but also for additional limbs [12, 43, 52]. For instance the supernumerary hand illusion is achieved simply by not hiding the real hand in the RHI, so that the participant sees both the real and the rubber hands being stroked. Furthermore, self- attribution of an additional limb does not appear to be limited to realistic body parts, but has for instance also been shown to apply for tails [119]. Bodily self-attribution can also happen with limbs or bodies of wrong size [11]. In a further example, body ownership was reported of a virtual arm that was slowly passively stretched to 4 times the normal arm length [64]. In my own experiment presented in LongArm (chapter 8) I could confirm that this also works in augmented reality (AR), when the arm extension is actively controlled by the user.

It has generally been suggested that we are more likely to feel body ownership of an artificial hand, if it looks realistic and is very similar to our own [6, 73]. Based on results from LongArm and PhysioMotor I can confirm this finding. However, it appears that the subjective similarity may have no such effect and the false assumption comes from the phenomenon that the BOI simply makes us perceive the artificial hand that we embody as more similar to our own [78]. In short: the body ownership illusion affects our perception of similarity and not vice versa [125].

In most VHI experiments the participants are instructed to focus on their virtual hand, which either remains motionless, or is moved slowly and in a controlled fashion. However, the VHI has also been successfully elicited during
interaction, while the participants’ foci do not lie on the virtual hand itself, but on some object they mean to manipulate [6, 73].

**Limitations of the body ownership illusion**

Botvinick and Cohen [15] found that the illusion does not occur with asynchronous multisensory stimulation, i.e., when the strokes are performed asynchronously on the real and the fake hand - a finding that has been confirmed many times since. Some researchers have also found that to elicit the BOI of a body-external object, it must resemble a plausible body part [96, 128]. For instance, Tsakiris and Haggard [127] showed that orientation and identity of the artificial hand are of importance, since a right rubber hand stimulated along with the subject’s left hand did not lead to the RHI. They also found that the rubber could not be replaced with any neutral object such as a stick. Costantini & Haggard [25] further evaluated the sensitivity of the RHI to spatial mismatches in posture and stimulation and found that a slight mismatch in the posture of the real hand is tolerated, while a mismatch in the posture of the rubber hand is not. Slater et al. [115] demonstrated that the RHI occurs under the condition that the arm is perceived as connected to the body (i.e., there is no gap). This was later confirmed by Tieri et al. [123], who studied the VHI for virtual hand representations that featured varying degrees of separation between hand and arm.

Ehrsson & Kalckert explored what they called the ‘Spatial distance rule’ [60] for the RHI by positioning the rubber hand at a vertical displacement to the real one (12cm, 27.5cm and 43cm). The authors found that the sense of body ownership weakens with distance: no ownership of the rubber hand was reported at 43cm displacement, and only sometimes at 27.5cm. A similar study by Lloyd [75] explored referral of touch when offsetting the rubber hand on the horizontal axis, reporting that the illusion significantly decreased from 17.5cm to 27.5 cm. However, it is apparent from the data that ratings remained positive for all hand positions. So it appears that the synchronous tactile feedback was enough to elicit a referral of touch illusion despite extreme displacement of the artificial hand. It must however be noted that ratings for referred touch are “more likely to evoke affirmative responses” [75] compared to ratings for body ownership. In a VHI experiment Nierula et al. [92] report that a 30cm offset between the real and virtual hand disrupts the analgesic effect, which has been found to be associated with the BOI. However, from the data it is evident that body ownership ratings remained positive with synchronous tactile feedback, despite lack of collocation. Furthermore, the authors find that an offset between the real and the virtual hands leads to a stronger illusion of owning more than one right hand.

With related work reporting contradicting results, it remains to be explored how robust the moving VHI is towards large vertical position offsets and how
2.3. FROM REALITY TO VIRTUALITY

this could be leveraged to improve interactions. Ownershift (chapter 9) takes a step towards answering these questions.

To summarize, according to the evidence presented above, some degree of visual, anatomical, volumetric, and postural congruency, as well as a certain spatial relation between the real and artificial hand, are critical for BOIs \[125\]. Based on the these requirements for BOIs it is likely that the sensory body representation (bottom-up modulation) is matched to a reference model of our body (top-down modulation), which is formed based on prior experience and innate representations of our anatomical and structural features.

2.3 From reality to virtuality

That which exists objectively is the reality we live in. What we perceive of this reality is what we think of as real; also called the subjective reality. Virtual reality (VR) is a digital, computer-generated (objective) reality that we perceive through visual, sound, or haptic rendering. VR aims to be so convincing that we can no longer distinguish it from reality, i.e., the interface should become completely unnoticeable.

In this thesis VR refers mainly to visual component. Most VR is visually experienced through one or more displays, which can be fixed (e.g., computer screen, CAVE), hand-held (e.g., smartphone), or head-worn (e.g., head mounted displays). Head mounted displays (HMDs) can either be monocular (one image for a single eye), bi-ocular (the same image for both eyes), or binocular (a different image for each eye) \[57, p.121\] \[107, p.42\]. There are three types of HMDs: non-see-through HMDs (e.g., Oculus Rift\(^3\)), video-see-through HMDs (e.g., the headset described in chapter 4), and optical-see-through HMDs (e.g., Microsoft Hololens\(^4\)) \[57, p.34\].

When VR content is combined with information from the real world, this is referred to quite generally as mixed reality. The virtuality continuum by Milgram and Kishino \[85\] maps systems based on their proportion of real visual content to virtual content (shown in Fig 2.2). VR is completely computer-generated and does not include any visual content from the real world. This type of experience is usually presented with a non-see-through HMD, or CAVE system. Augmented virtuality brings real world content into VR (e.g., with a video-see-through HMD). However, the visual elements are primarily virtual. Augmented reality (AR) on the other hand adds virtual content to the real world (ideally in a way so that we can’t distinguish what’s real and what isn’t). This can for instance be achieved with an optical-see-through HMD or video-see-through, where most visual elements are part of the real world, and only a small portion of virtual content is added.

\(^3\)Oculus Rift: https://www.oculus.com/rift (last accessed: 28.06.2018)

CHAPTER 2. BACKGROUND

The precise placement of a system on this continuum is difficult, since it seems impossible to accurately quantify the amount of virtual vs. real content. I refer to the system presented in chapter 4 as an AR system, since the user mainly perceives the real world around him with only a minor portion of added virtual content (the virtual arm). The system presented in chapter 5 is a VR system, since the user perceives only the virtual environment around him.

Depth cues in VR and AR

Depth cues are specially important in our personal space, since this is the area within our arm’s reach (or slightly beyond) in which we naturally interact with the world. In this space our binocular vision and proprioceptive cues help us interact. Beyond this space lies the action space (at about 2m-20m), in which we can access things quickly by moving around. Even further away lies the vista space, which is mostly beyond our control [57, p.112].

If well calibrated, binocular HMDs lead to stereopsis (stereo-vision), but can also cause unease due to conflicting vergence and accommodation of the eyes. Furthermore, binocular depth cues are strongly diminished beyond our personal space and it appears that 3-5% of people are stereo-blind. Generally, the benefit of binocular over bi-ocular vision remains debated [57, p.121] and the implementation of additional depth cues in VR and AR is crucial.

The most effective depth cue across all distances is occlusion, which simply means that closer opaque objects will hide objects that are further away when they overlap [57, p.114]. When occluding virtual objects with objects from the real world in AR, so-called ‘phantoms’ are used [107, p.199]. Shadows also play an important role in indicating the location of objects, thereby providing information about depth [57, p.116] [107, p.220]. Apart from these depth cues, which are discussed in more detail in chapter 4, there are many others, such
2.4. CONCEPTS FOR INTERACTION IN VIRTUAL REALITY

as relative size, linear perspective, texture gradient, and motion depth cues.

Immersion and presence

Immersion refers to the capability of technology to present virtual content in an inclusive (shutting out reality), extensive (multisensory), surrounding (wide field of view) and vivid (high resolution, richness of information) fashion. Furthermore, “immersion requires a self-representation” [114] in the virtual world - in other words, a virtual body. In this context it is important that the posture and movements of the virtual body match to our own proprioceptive and motor signals. The virtual body forms part of the virtual environment, represents the actor and perceptor, and is at the center of the virtual world.

Immersion is subjectively experienced by the user as presence. Presence is the feeling of ‘being there’, which makes the virtual environment feel like a place one visits. Presence does not require photorealism, or a perfect reconstruction of reality. It only requires the critical stimuli by which our brain can then “fill in the gaps” [57, p.61]. Presence does however depend on a responsive system, depth-cues, and matching motion of the avatar. Also, cues from the real world, such as the experimenter’s voice or low screen resolution, can remind the users of where they actually are and thereby cause a break-in-presence [114].

Presence is useful to achieve, since it makes us behave like we would in the physical world. This is particularly relevant for training applications, psycho therapy, and research on cognition.

2.4 Concepts for interaction in virtual reality

Despite arguably being at the very core of HCI, interaction is a vague and under-defined term with many different interpretations. Hornbæk and Oulasvirta identify seven distinct concepts of interaction [51]. I understand interaction as a mix of several of these concepts. Interaction-as-tool-use regards the computer system as a tool that is used to achieve a goal. The tool or interface mediates between the users and the domain object they want to manipulate. A good tool or interface should be useful and unnoticeable, or even transparent, by affording direct manipulation. It could also amplify the users’ abilities, for instance by extending their reach. However, I interpret human reaching movement in terms of interaction-as-control: the human generates input (e.g., lifting a hand) that is modulated based on output (e.g., eye-hand coordination), to achieve a minimization of error towards a certain goal (e.g., reaching for something). To evaluate interaction I apply measures from the concept interaction-as-transmission (e.g., Fitts’ Law), among others. Furthermore, in interaction-as-tool-use the human and the computer are clearly separate entities, whereas in interaction-as-embodiment no such distinction is made. While I started out designing systems regarding the human and the system as
CHAPTER 2. BACKGROUND

Figure 2.3: **Combining the physical and the digital world.** As a step beyond command languages, direct manipulation interfaces allow users to directly manipulate a representation of data. (A) In the first generation, with Graphical User Interfaces (GUIs), the screen provided the border between the physical and digital worlds. (B) Tangible User Interfaces (TUIs) and (C) Radical Atoms [55] provide a physical representation of the digital world, thereby moving this border into the real world. (D) Extending the body through the body ownership illusion goes beyond these paradigms by moving the border between the physical and the digital into the user’s own mind.

separate units, it has become the goal of my work to dissolve this boundary between ourselves and the computer system, between real and virtual, in order to feel connected to the virtual hand representation through which we interact (Fig 2.3). Thus it becomes obvious that in practice no clear categorization into such concepts is possible.

**Direct manipulation**

The term *direct manipulation* was introduced by Shneiderman in the 80’s [112] as the feeling of directly manipulating an object of interest, instantly perceiving the changes and allowing immediate reactions and corrections by the user, versus communicating with an interface through commands which are difficult to learn and receiving unexpected or delayed feedback, such as error messages. Hutchins et al. [53] mention *distance* and *engagement* as specific factors for direct manipulation. Engagement describes the feeling of directly manipulating the object of interest. The term distance describes the figurative gap between the user’s objective and the action that completes the task. This distance should be small to reduce cognitive overhead, since the feeling of directness is inversely proportional to cognitive load. Later Beaudouin-Lafon [13] redefined direct manipulation for Post-WIMP (windows, icons, menus, pointer) interfaces which are zoomable, include interactive visualizations instead of icons, and widgets instead of menus, and are even more natural and direct. With the interaction techniques presented in LongArm (chapter 8) and Ownershift (chapter 9) we achieve direct manipulation in a sense that is perhaps closest to the definition of Beaudouin-Lafon. However, one could argue that, with
immersive VR, direct manipulation has reached a whole new level, where 3D interaction techniques allow the object of interest to simply be grabbed and manipulated like physical objects in the real world.

3D interaction techniques

Interaction techniques in 3D often involve “gestures in free (unconstrained) 3D space” [5] which are frequently divided into two phases: selection and manipulation. Selection typically precedes the manipulation task and existing approaches can roughly be categorized into two groups, which we find described as virtual hand (e.g., 3D cursor) and virtual pointing (e.g., ray casting). While virtual pointing is reported to be more effective, more accurate and less strenuous, virtual hand allows a direct transition to the manipulation task. I prefer to focus on the virtual hand approach since there is no separation of selection and manipulation tasks when we use our hands to interact with the real world. An example for 3D interaction with a virtual hand is the Go-Go Interaction Technique by Poupyrev et al. [99], where users can reach out with an abstract hand to select distant objects in immersive VR.

Bowman et al. [16] provide a useful set of guidelines for creating 3D interaction techniques with good usability. For instance, there should be no interpenetrating objects in VR, since the solidity of real-world objects prevents them from it and thus makes it feel wrong. The authors also recommend occlusion and perspective as depth cues for interaction with distant objects. These guidelines were particularly relevant for the interaction technique presented in LongArm, where I addressed these aspects by implementing collision detection, occlusion effects, and shadows of the virtual arm on the virtual representations of real-world objects (discussed in more detail in chapter 4.3). Furthermore, interaction should be made possible only with visible objects. This can lead to a problem of space management, since occlusion, clutter, and distance make accurate target selection and manipulation difficult. These factors might make some kind of navigation necessary, as can for instance be achieved through the distant reaching techniques discussed below.

Distant reaching

The need for extending the user’s reach frequently arises in the context of interactive tables or large vertical displays. For instance, users can use a virtual hand extension called I-Grabber [1] to reach remote objects on a collaborative interactive tabletop. The I-Grabber is created by placing both index fingers on the table next to each other and moving them 20cm apart. Moving the fingers further apart then extends the grabber’s arm, allowing selection remote objects. Parker et al. presented a hybrid point-touch input technique called TractorBeam [94], by augmenting a stylus to allow remote pointing (laser-style pointing) on an interactive touch table. Further, Vogel et al. [131] present
three techniques for gestural pointing and clicking on large planar displays, where a ray is cast out from the user’s finger.

Similar approaches exist for distant pointing and selecting in VR, an overview of which is conveniently given in a survey by Argelaguet and Andujar [5]. For instance, the Go-Go Interaction Technique [99] mentioned earlier is a distant reaching technique where the user’s arm length can be dynamically increased to reach vases placed on distant pedestals. The user’s physical space is hereby divided into two zones, the closer zone where the hand position is directly mapped to the virtual hand position, and the outer zone where the arm length increases non-linearly. This way users can seamlessly interact with both near and far objects. This technique largely inspired my implementation of the LongArm-technique (chapter 8). Similarly, FingARtips [18] offers natural fingertip-based manipulation of distant virtual objects in AR, albeit with the use of a claw instead of a hand-shaped cursor.

Virtual space transformations

VR affords control of many factors, such as the appearance of objects, the feedback users receive, how fast they can move, and also the extent of space. The Go-Go Interaction Technique [99] described earlier, is an example for a transformation of space in VR, where users could actively extend their virtual arm, to bring remote virtual objects in the room ‘closer’. But the virtual space can also be shifted completely without users noticing, as has been shown with redirected walking [101]: While visual cues tell the users that they are traversing a large virtual room, in reality they are repeatedly walking back and forth on the same arc in a much smaller space. This allows exploration of a large virtual environment through walking in a much smaller tracking area, and enables the use of a single haptic proxy for multiple virtual objects [67].

Such redirections are not only limited to walking. They can also be applied to the user’s hand during simple reaching tasks. For instance Haptic Retargeting [8] describes re-using a single haptic proxy for multiple virtual targets by ‘tricking’ the human perception. While participants perform pick and place operations with multiple virtual cubes, a slight incremental rotation of the virtual body or the virtual world redirects their hand so that they always reach to pick up the same physical cube. The authors compare three methods for warping the users’ space: in the Body Warping approach they slightly rotate the users’ virtual body during the reaching movement, leading them to correct their movement to reach the target. Secondly, in the World Warping approach the world is slightly rotated whenever the users turn their head, to achieve alignment of the virtual target with the haptic target. The third method, Hybrid Warping, is a combination of both previous approaches, so that each accounts for 50% of the warp. The virtual targets are placed with offsets requiring horizontal rotations of 16°. Furthermore, the Body Warp and Hybrid
2.4. CONCEPTS FOR INTERACTION IN VIRTUAL REALITY

Warp both allow vertical warping as well, which was used in a cube stacking task. The study shows that use of a haptic proxy improves both presence and satisfaction, and that Hybrid Warping proved most successful. Another retargeting approach by Montano et al. [86] aims to make controls in VR more easily accessible rendering interaction less fatiguing. The authors describe an implementation similar to Body Warping, albeit as a translation of the hand towards the targets. This technique allows interaction with multiple targets in an open-ended fashion (the warp only depends on the hand’s position - no knowledge of the target is required). To achieve this the physical and virtual space are both divided into corresponding tetrahedrons, which differ slightly in shape. Each point in a physical tetrahedron corresponds to one point in a visual tetrahedron (isomorph), thus when the users move their hand into the physical space of such a pyramid, the virtual hand is rendered at the corresponding position in the visual pyramid. The authors pursue two approaches: 1) Spatially consistent and 2) Ergonomic. In the first the distances between targets remain similar, whereas in the latter these rules are loosened to allow more ergonomic optimization.

Most techniques discussed above used only small shifts (e.g., max. 10 cm [86]), since they aimed to make the transformations unnoticeable to the user. In exploring larger offsets for haptic retargeting in VR, it was found that up to $40^\circ$ deviation between the paths of the real and virtual hand were perceived as “bearable” [23]. A different experiment explored horizontal offsets of up to 76 cm between the real and the virtual hand [45], and found that we can quickly adjust to an instantly shifted virtual hand space and reach better performance than with large interpolated shifts. Performance measures for interaction are discussed further in chapter 6.

**Ergonomics of interaction**

Interactions that involve large gestures with a raised arm, such as when interacting with large screen displays, have been found to lead to fatigue, which is colloquially known as the ‘Gorilla Arm syndrome’\(^5\). This has sparked a lot of research on fatigue during interaction with large touch screens [3], along with interaction techniques for reducing strain [74]. However, the ergonomics of interaction are important not only in context of large vertical touch screens, but for a much wider range of interfaces, including smartphones and laptops [9], as well as mid-air interaction in VR [10, 86].

Multiple models exist that help identify the optimal regions for interaction (or action) and allow us to design low-fatigue interfaces. As an example, the RULA (Rapid Upper Limb Assessment) system [82] was developed to quickly evaluate repetitive movements that industrial workers needed to perform and

\(^5\)Gorilla Arm syndrome: https://www.computer-dictionary-online.org/definitions-g/gorilla-arm.html (last accessed: 28.06.2018)
identify possible health risks due to fatigue and bad posture. This system allows scoring of different postures and movements, based on the orientation of the limbs and upper body in respect to each other. According to RULA, the least fatiguing posture for prolonged manipulation includes a straight upper body, relaxed upper arms and slightly bent elbows, with the hands in front of the body at waist level. Montano et al. [86] applied a variant of RULA to evaluate their retargeting approach described previously. Hincapié-Ramos et al. [47] devised a different model to evaluate shoulder fatigue, or as they call it ‘consumed endurance’, where the human arm is modeled as a single joint. The authors claim that unsupported mid-air interactions are least strenuous when the hands are located about mid-way between the waist and shoulders, allowing slightly bent elbows. Using biomechanical simulation, Bachynskyi et al. [10] explored a wide range of arm movements in reachable space with a clustering approach, which is based on muscle activation and performance. Results show lowest muscle activations for short to medium length movements in the lower near part of reachable interaction space. In summary, the models discussed above agree that the optimal placement for interactive content is approximately at the user’s waist level and within comfortable reach, so that the arms do not need to be raised or extended over a long period of time.

2.5 Positioning of own work

To provide the reader with a clear idea of my contributions to each field of research, I will shortly position each of the papers attached in Part II (LongArm [34]: chapter 8, Ownershift [35]: chapter 9, PhysioMotor [36]: chapter 10) in the context of relevant related work.

Interaction techniques are typically evaluated based on task performance (e.g., accuracy, task completion time). I propose body ownership as another measure by which to evaluate interaction, where we focus on the users’ experience instead. This can be used to evaluate distant reaching techniques [1, 18, 39, 94, 99, 131], as presented in the LongArm publication. Similarly, we can evaluate techniques that ease overhead interaction based on the experience of body ownership of the virtual hand (Ownershift).

In regards to redirection techniques in VR, I have found that most related systems apply only small horizontal shifts, either for the purpose of guiding the user’s hand to haptic props [8, 23], or to reduce fatigue [86]. However, I argue that larger offsets are required to effectively reduce strain of overhead interaction, which I explored in Ownershift. What further sets my work apart, is that the proposed Ownershift technique allows fast reaching towards the target with 1:1 mapping of hand positions, and the hand space is only gradually transformed during prolonged interaction (as is also described in chapter 5). I am not aware of any previous work exploring shifts that are gradually applied
2.5. POSITIONING OF OWN WORK

over time, as opposed to an instant shift [45], or interpolation based on the location of the real hand [8, 23, 86, 99].

In regards to BOIs, I find that the vast majority of VHI studies are implemented in VR. I find it relevant to explore body ownership in AR, since this would arguably allow more meaningful interactions with physical objects in our real environment. Here a core difficulty lies in bridging the gap between virtuality and reality in a convincing fashion. The AR system presented in chapter 4 describes an approach for tackling such challenges. LongArm presents the VHI experiment based on this system.

My publications further contribute new insights as to the flexibility of BOIs. For instance, it has been shown that the BOI can be preserved when the virtual arm is passively stretched [64], but it remained unclear whether the illusion would also persist if this transformation of the virtual hand was controlled by the users themselves (LongArm). Furthermore, mismatches in location of the real and virtual hand were so far reported as detrimental for BOIs [60, 75, 92]. The Ownershift experiment shows that even large vertical position offsets can in fact be tolerated. Last but not least, PhysioMotor contributes an extensive user study on how different multisensory stimuli (i.e., visual-proprioceptive, visual-motor and cardio-visual stimuli) affect the VHI for unrealistic hand representations. The findings show that the type of multisensory stimuli is not critical for realistic virtual hands. However, for more unrealistic hand representations, visuo-motor stimulation is most powerful in eliciting a BOI and may even allow to compensate for the lack of limb-connectivity.
System design and implementation

The computer systems targeted by my research provide an immersive virtual or virtually augmented environment, which the users experience through a head mounted display (HMD), and with which they interact using nothing but their own hands. A network of sensors and actuators, either in the environment or wearable, keep track of the users, enable interaction with the virtual (or virtually augmented) environment, and provide feedback.

Virtual reality (VR) is a powerful tool for exploring the body ownership illusion (BOI) since it provides a high degree of control over the environment and the user’s body representation. I take advantage of this technology, not only to further understand the nature of BOIs, but to explore their application in interactive systems.

Over the course of my PhD I developed two such systems, which served as platforms for experimentation. Chapter 4 presents details about the system that was used to explore the virtual hand illusion (VHI) in augmented reality (AR) and on which LongArm [34] (chapter 8) is based. Chapter 5 describes the immersive VR system, which was the platform for the research presented in Ownershift [35] (chapter 9) and PhysioMotor [36] (chapter 10).

According to Spanlang et al. [118], the required technical infrastructure for embodiment of a virtual user representation consists of the following components: an HMD, head tracking, motion capture, and a module for multisensory

Figure 3.1: System components. General components of a virtual embodiment system.
feedback (e.g., vision and haptics). Hence, the main components of both my AR and VR systems are the **Display module** and the **Tracking module**, as depicted in the diagram in Fig 3.1. The Tracking module is of special importance because direct manipulation of virtual objects in unconstrained 3D space requires capturing of the user’s hand movements and transferal of these movements to the virtual user representation (i.e., the virtual arm).

Apart from the HMD and motion tracking technology, each of the described systems included additional hardware. These components, listed under **Other** in Fig 3.1, were added based on requirements of the individual interaction technique or experimental design. The AR setup used for LongArmincluded an Actuator module, allowing the user to control actuated furniture in the room. For Ownershift I included a Haptic module, which provided vibration feedback during interaction with virtual objects. PhysioMotor featured a Physio module, permitting visualization of the participant’s heartbeat on the virtual hand.

The experiments and interaction techniques discussed in this thesis were mainly implemented in the Unity\(^1\) game engine. Further details about the implementation and the system components are given in the following chapters (AR system: chapter 4, VR system: chapter 5). These may serve to inform the design of future immersive VR and AR systems that aim to provide an embodied experience in an immersive virtual, or virtually augmented, environment.

### 3.1 Iterative design and development

Each of the systems described above was developed and evaluated iteratively, involving multiple pilot studies. These pilot studies were usually quite short with only few participants (for specific numbers please refer to chapter 6.1), whom I often recruited from among my work colleagues. Pilot studies served multiple purposes, depending on the state of development: Firstly, they provided a form of quality assurance and allowed me to verify the overall functionality of the systems, from the user calibration to the interaction technique. At this point the system was often still incomplete and I focused on exploring just one or two features at a time. Secondly, these studies served to inform design choices and identify suitable parameters. Examples are scaling parameters for hand transformations, placement of interactive targets, the configuration of tracking sensors, design of multisensory feedback, and the choice of certain hand visualizations. The following sections will provide more details on all of these aspects.

---

\(^1\)Unity Game Engine: [https://unity3d.com](https://unity3d.com) (last accessed: 28.06.2018)
3.2 Designing for body ownership

Multisensory stimuli

The key to eliciting the illusion of owning a virtual hand are congruent multisensory stimuli. According to related work, there are multiple possibilities: from the classic stroking of fingers for visuo-tactile stimulation, to integration of visuo-motor stimuli when a virtual hand moves like the own, to so-called ‘vision-only’ approaches that involve integration of vision and proprioception and lead to body ownership simply by looking at the artificial hand. A slightly different idea involves the use of interoceptive signals (e.g., heartbeat), which can for instance be visualized on a virtual body representation, thereby potentially strengthening the user’s connection to the artificial body.

In my experiments I applied and tested several combinations of sensory input to support the illusion of owning a virtual hand representation through multisensory integration. These are discussed in more detail in chapter 4.2 (AR system) and chapter 5.2 (VR system).

Virtual user representation

Since it has been shown that the appearance of the virtual hand is critical for the ownership illusion [96, 127, 128], an important part of the iterative design was the evaluation of different virtual hand representations. To trick the brain into perceiving a virtual hand representation as belonging to the own body, this virtual hand should to some degree resemble a human hand (e.g., in morphology, size, color), without however falling into the uncanny valley [87]. It is yet unknown what level of realism is required for the illusion to occur and this may also be subject to strong interpersonal differences.

More details on the evaluation of virtual hand representations are given in context to the AR system in chapter 4.2 and the VR system in chapter 5.2.

To summarize, I chose realistic-looking hand representations to reliably elicit the BOI and explored reactions to less realistic virtual hands in contrast to these. My findings confirm that realistic hand representations are beneficial for the BOI. I have also found that lacking connectivity between the user’s body and the hand, or between the fingers and the wrist, was detrimental for the BOI of the virtual hand, as was previously found by Tieri et al. [123]. However, this limitation was less strong when visuo-motor synchrony was given, i.e., when observing natural hand movement of a disconnected virtual hand, our mind may fill in the gaps.

User calibration and scale of the virtual environment

For a virtual user representation to feel most acceptable, it needs to be adjusted to the user. For instance the height of the virtual body and the elevation
CHAPTER 3. SYSTEM DESIGN AND IMPLEMENTATION

of the camera in the scene should be similar to our own body height and elevation of our eyes. We notice quite strongly if the camera is too low, since the ground seems too near, and if it’s too high we suddenly feel taller than we are used to. Also, the length of the virtual arm and the size and appearance of the virtual hand should ideally be similar to our own, to improve the experience. The systems presented in this thesis account for the first by calibrating the avatar’s height and position of the camera based on the user’s height, which was either indicated by the users themselves, or gained from tracking data of the HMD or external tracking systems. The aspect of the arm length was specially important for extending the virtual arm in LongArm (chapter 8). To achieve a good match the participants were asked to extend their right arm. Then the 3D model was automatically rescaled, to match the their own arm length (measured from the shoulder marker to the hand marker). For Own-ershift and PhysioMotor, which are based on the VR system (chapter 5), the hand tracking sensor handled adaptation of arm length and hand size automatically. In regard to the appearance of the virtual hand, efforts were made to use hand models that could fit both genders and reflected the users’ apparel (e.g., when users were required to wear gloves). Approaches for dynamically creating textured 3D models from camera footage could help to provide more realistic virtual avatars in future.

Apart from the virtual user representation, also the environment should have a realistic scale to create a strong feeling of presence. This is particularly important, if it contains objects from everyday life that we are very familiar with. For instance, the furniture in a room should appear to have realistic proportions, doors should be high enough to walk through, etc. This can be challenging to adapt, since VR provides no fixed scale (i.e., you can define yourself how long 1 meter is and scale everything according to that). Both systems presented in this thesis, were built in Unity² and frequently the use of ‘1-meter’ blocks proved helpful as a reference for scaling virtual objects to a realistic size.

²Unity Game Engine: https://unity3d.com (last accessed: 28.06.2018)
Augmented reality system for body ownership

The system presented in this chapter was developed with the aim to study the virtual hand illusion (VHI) in AR during interaction with the physical environment. To achieve this, the users’ virtual arm is extended, allowing access to remote physical objects through AR. Thus, users can directly manipulate objects from a distance using their bare hands. The results of this are presented in LongArm [34] (chapter 8), where we explored whether we could make users feel like a long virtual arm was actually part of their own body, while they were using that arm to control remote actuated furniture in the room. In this experiment we compared three hand representations that differed in their degree of realism and visible connectivity to the user’s body, and we varied the visibility of the user’s own arm, exploring the effect of these factors on the VHI.

4.1 AR system components

The AR system consists of three main components, depicted in Fig 4.1: the Display module, the Tracking module and the Actuator module. These are described in more detail in the following sections.

Figure 4.1: Components of the AR system. The AR system consisted of a Display module with a video-see-through HMD and a Tracking module relying on OptiTrack. The Actuator module facilitated interaction with actuated components (i.e., furniture) in the environment.
CHAPTER 4. AR SYSTEM FOR BODY OWNERSHIP

Figure 4.2: Setup for the LongArm experiment. (A) The user could control actuated furniture in the room, such as a rotating surface (white table), a large desk (wooden table), and a double curtain. (B) Modified Oculus Rift DK2 with front-mounted camera. (C) Retro-reflective tracking markers were worn by the user for motion tracking. (D) The user’s view in AR.

The Display module for video-see through AR

This system relied on an Oculus Rift DK2\textsuperscript{1} with a stereo-camera mounted on the front to provide video footage of the user’s real surroundings (see Fig 4.2, B). The camera was produced by the startup company Ovrvision\textsuperscript{2} and could theoretically supply stereo-vision. However, due to difficulties with the driver software, only one of the cameras was operational. While this means that I could not take full advantage of the stereo-camera, seeing the bi-ocular camera footage (single camera) through the binocular head mounted display (HMD) at least provided stereo augmentation [107, p.43] and did not hamper the functionality of the system.

By displaying the live video feed in the Oculus, showing the user a view of the real world, the headset was effectively turned into a video-see-through HMD. To achieve accurate mapping and compensate for the decentral placement

\textsuperscript{1}Oculus Rift DK2: https://www3.oculus.com/en-us/dk2 (last accessed: 28.06.2018)
\textsuperscript{2}Ovrvision: http://ovrvision.com (last accessed: 28.06.2018)
4.1. AR SYSTEM COMPONENTS

Figure 4.3: Tracking module. The OptiTrack system allows accurate motion tracking of the user and objects in the scene. (A) shows a view from one of the Flex13 Optitrack cameras depicted in (B). (C) shows the markers that the participants were wearing during the LongArm experiments.

of the camera, the camera coordinate system was manually aligned with the tracking space. This was done by creating a distribution grid for the camera’s field of view, based on the instructions of Steptoe\(^3\). The virtual camera was then adjusted to match the perspective properties of the physical camera.

Motion tracking module

To allow direct manipulation of virtual objects in unconstrained 3D space it was important to track the user’s body and hands, and transfer these movements to the virtual user representation (in particular the virtual arm). Therefore the system was deployed in a 4x4 meter tracking area with 10 Flex13\(^4\) cameras (Fig 4.3, B) mounted on surrounding trusses. Full-body motion tracking was supported through the OptiTrack Motive\(^5\) software (Tracker package, version 1.10). This software allows accurate tracking of objects with 6 degrees of freedom, by attaching a set of retro-reflective markers to form so-called rigidbodies.

For the LongArm experiment several retro-reflective markers were attached to the front and sides of the HMD to track head movements. A unique set of markers was also attached to each shoulder to allow tracking of the user’s body orientation. For tracking of arm movements, markers were attached to


CHAPTER 4. AR SYSTEM FOR BODY OWNERSHIP

Figure 4.4: Actuator module. The Actuator module consisted of a motorized, movable component, e.g., a Linak table leg, which is controlled through an Arduino Uno.

the right upper arm close to the elbow, and on the back of both hands (see Fig 4.3). The OptiTrack system was run on a separate PC and streamed to the PC running Unity in real-time using the Natnet SDK.

An advantage of OptiTrack, is that it is highly accurate (according to the manufacturers markers can be tracked with sub-millimeter accuracy) and provides reliable outside-in tracking, allowing the user to move freely within the entire tracked volume. The main weakness of this tracking method is that the system is very costly, requires fixed installation, and it requires users to wear markers. Furthermore, external infrared sources (e.g., sunlight) cause disturbances, it cannot deal with occlusions, the number of trackable rigidbodies is limited, and last but not least, tracking of individual finger movements was not (and to my knowledge still is not) supported.

Actuator module for room control

In the Actuator module Arduino Uno\(^7\) microcontroller boards provide access to external electronic components, such as actuated furniture. These are programmed using the Arduino IDE, to receive commands from Unity via serial communication and send signals to the electronic components accordingly. For LongArm three such Arduinos were used to control electric motors, which moved two tables and a curtain shown in Fig 4.2 (A). The first table was a


\(^7\)Arduino Uno: https://store.arduino.cc/arduino-uno-rev3 (last accessed: 28.06.2018)
height-adjustable office desk with motorized legs of the brand Linak\textsuperscript{8}, which allows precise height-adjustment. The second table was a self-constructed contraption, which featured a single motorized Linak lifting column that allowed rotation of the tabletop from horizontal to an angle of about 45°. Each of these tables was originally controlled by pressing two buttons - one for ‘up’ and the other for ‘down’ (Fig 4.4, top-left). These buttons were replaced with a relay and an Arduino that would simply close the ‘up’ circuit, the ‘down’ circuit, or neither, based on the commands it received from Unity. The electronic curtain could originally be controlled with 3 buttons (open, close and stop). These were again replaced by a relay and an Arduino, which closed circuits depending on the user’s input to trigger the respective action.

4.2 Elicitation of body ownership in AR

Multisensory stimuli

LongArm relied partly on visuo-proprioceptive integration, since the real and virtual arm were collocated when close to the body (1:1 mapping). (See section 4.3 for details on the implementation of the arm-stretch technique.) Furthermore, the users experienced synchronous visuo-motor feedback both in 1:1 mapping and for the stretched virtual arm, since the virtual hand moved according to their own hand and they could actively control the amount of arm stretch. A limitation was the lacking support of finger tracking. The manipulation of the interactive elements in the room did not require finger movement for elaborate gestures, and in the experiments participants were instructed to keep their hand rigid with loosely extended fingers. However, should a user choose to wriggle the fingers or make a fist, the illusion of ownership would most likely break.

Virtual user representation

In the LongArm experiment we explored the effect of visual appearance on the body ownership illusion, by comparing three hand representations shown in Figure 4.5. For this purpose I intentionally designed a less realistic hand representation, consisting of a simple extruded hand shape. This abstract-hand was compared to a more realistic-looking virtual hand that was either disconnected (which in the paper is referred to as hand), or connected to the virtual body through an arm (referred to as arm). Care was taken to give all hands the same general look (i.e., colors), so that they only differed in their morphological complexity and connectivity.

\textsuperscript{8}Linak lifting columns: \url{https://www.linak.com/products/lifting-columns} (last accessed: 28.06.2018)
CHAPTER 4. AR SYSTEM FOR BODY OWNERSHIP

The most realistic hand (Fig 4.5, A and B) was part of an avatar called ‘Ethan’ that was contained in the standard assets package in Unity v.5.1. This hand fit very well for the purpose, since the participants needed to wear fingerless black gloves (with attached retro-reflective markers for motion tracking), which were also part of the avatar’s costume. For the purpose of the experiment, the texture of the avatar was modified so that the arm and hand resembled that of the participant (skin colored arm and black glove), while the rest of the avatar’s body was rendered invisible.

The abstract-hand was designed in multiple iterations. In the first design iteration it consisted of a skin-colored cone on a short rod (Figure 4.6, A), resembling a 3-dimensional mouse cursor. In a second iteration the cone was replaced with an oblate spheroid, which resembled more the flat of the hand and helped indicate the rotation of the user’s palm (Figure 4.6, B). Then followed a third variation, which was inspired by the type of hand cursors that are often used for mouse interaction (e.g., the hand tool that is used to drag pages up and down in the Adobe\textsuperscript{9} pdf reader). This extruded hand shape (depicted in Figure 4.5, C, and Figure 4.6, C) created a stronger ‘hand’ association and also allowed a more direct comparison to earlier work (Poupyrev et al. [99]).

When interacting with their environment using the arm representation (Fig 4.5, A), it appeared to the participants as if the virtual arm was connected to their own body at their right shoulder. Since this system did not support finger tracking (the virtual hand moved according to the user’s hand, but finger movements were not supported), participants were asked to keep their hand open with their fingers loosely extended. This also made the realistic hand (arm and hand representations) comparable to the abstract-hand.

\textsuperscript{9}Adobe Reader: https://get.adobe.com/reader (last accessed: 28.06.2018)

Figure 4.5: Virtual user representation. Three virtual hand representations were compared in the LongArm experiment: (A) arm: a realistic-looking, connected hand, (B) hand: a realistic but disconnected hand, and (C) abstract-hand: a simplified hand representation.
4.3 Implementation details

This section describes the implementation of the interaction technique and the creation of the virtually augmented environment.

Interaction technique

The interaction technique is based on the Go-Go-Interaction Technique [99]: While the users’ hand is located close to their body, the virtual hand follows it directly with 1:1 mapping (Fig 4.7, A-B). However, once the users extend their arm beyond a certain threshold ($R_r > D$, with $R_r$ as the distance from physical hand to shoulder and $D$ as the threshold), the virtual arm is stretched (Fig 4.7, C-D). The non-linear arm-stretch occurs along a vector drawn from the users’ right shoulder to their real right hand, so that the virtual hand is located at a distance

$$R_v = R_r + k \times (R_r - D)^p$$

from the users’ body. The factors $k$ and $p$ define speed of length increase and the maximal arm length, and are discussed further below.

Defining the maximal arm length

The stretch function was adjusted and evaluated in various iterations, since controlling a long virtual arm is challenging for multiple reasons. Firstly, I found that depth perception was generally weak, which made it difficult to position the virtual hand exactly at a certain distance from the body. Our perception of space through an HMD is generally distorted or compressed.

Figure 4.6: Abstract hand representation. Several iterations were made for the design of the abstract hand representation, beginning with the arrow (A), which was a very basic composition of geometric shapes to roughly resemble a 3-dimensional arrow. This was replaced by the paddle (B), which allowed better judgment of orientation. Finally I created a simple, extruded hand-shape, the abstract-hand (C).
CHAPTER 4. AR SYSTEM FOR BODY OWNERSHIP

Figure 4.7: Arm-stretch technique. When holding the hand close to the body (A), the virtual hand follows the user’s hand with a 1:1 mapping (B). Moving the hand away from the body beyond a certain threshold (C) activates the arm stretch, which non-linearly increases the distance of the virtual hand to the user’s shoulder (D). The virtual arm is thereby stretched beyond the user’s natural arm length, effectively extending the user’s reach.

[57, p.124], which may have further been aggravated by the lack of stereo-vision and the distortion of the camera image. Furthermore, our visual acuity becomes lower at a distance. Secondly, extending the hand far from the body becomes physically straining, if such a pose is to be maintained for a prolonged period of time. Since the actuated furniture used in the LongArm experiment moved fairly slowly, the users were forced to perform slow motions.

Hence, I aimed to configure the arm-stretch function (see Fig 4.8) such that the most distant object in the room (the curtain at 2.2 meters away from the user) could be reached comfortably, without full extension of the arm. I found $p = 4$ to be a good value, and both $D$ and $k$ were calculated to achieve a maximum arm length of $R_v = 5$. That means that when the user’s real arm was fully extended, the virtual hand appeared to be 5 meters away from the
4.3. IMPLEMENTATION DETAILS

Figure 4.8: Arm-stretch function. When the user’s arm is extended beyond a threshold distance $D$, the virtual arm stretches, increasing its length nonlinearly.

user’s shoulder.

Direction of the arm stretch

Identifying the best direction for the arm-stretch was far from intuitive. In pilot studies I evaluated stretching in multiple directions, drawing a vector to the hand marker from either the head, the sternum, the shoulder, or the elbow. Stretching along a vector from head to hand results in an arm extension along the line of sight, which prior work has found useful for selection of distant targets [58, 134]. However, since the arm is attached to the body at the shoulder, this led to a strange orientation of the virtual arm, which did not match with the user’s real arm (see Fig 4.9, A). Stretching along the sternum-hand vector resulted in matching arm alignment in the user’s sagittal plane, but mismatch in the transversal plane (Fig 4.9, B). Stretching the arm along the vector drawn from the elbow to the hand causes the strongest misalignment between real and virtual hand and was very difficult to control (Fig 4.9, C).
CHAPTER 4. AR SYSTEM FOR BODY OWNERSHIP

Figure 4.9: Direction of arm-stretch. Design iterations of the arm stretch involved extending the virtual arm along (A) the head-hand vector, (B) the sternum-hand vector, (C) the elbow-hand vector, and (D) the shoulder-hand vector.

Finally, I found that drawing the vector from the shoulder to the hand feels most natural and intuitive (Fig 4.9, D).

Room layout and implementation

In the LongArm experiment, the users sat on a black stool in the center of the tracking area, as shown in Fig 4.3 (A) and Fig 4.2 (C). A round black table was placed next to them to support their left arm and keep their torso steady. The large motorized desk was located in front of the users, with the electric curtain mounted on the trusses just behind it. Next to the desk and to the users’ left was the flip-table (Fig 4.2 (A)). All actuated elements were placed out of the users’ reach, with the closest object at about 110cm away (front edge of large desk) from the user’s horizontal position, and the most distant object about 220cm away (curtain handle). OptiTrack markers were placed on all actuated elements in order to track their positions.

Apart from the non-linear stretching of the virtual arm, the virtual hand representation follows the user’s hand motions exactly. This allows users to directly and naturally interact with objects in the environment that are out of their physical reach. For instance, reaching out to the top of the desk with the virtual hand, the users could ‘push’ it down making it move down (see Fig 4.2 (D)), and likewise they could raise it again by reaching underneath pushing up. The flip-table behaved similarly: pushing down or up on the edge of the table top changed the incline of its surface. To allow control of the curtain a handle was fastened on one side. The user could then open and close the curtain by pushing this handle to the left or right.
4.3. IMPLEMENTATION DETAILS

Figure 4.10: Virtual object representations. The interactive objects in the room were modeled through simple bounding boxes in Unity (A), which allowed collision and clipping of the virtual hand (B), as well as visualization of virtual hand-shadows on the table top (C).

Implementation of depth cues

In Unity the actuated real-world objects are represented as transparent bounding boxes (e.g., the desk is visualized as large blue box in Fig 4.10, A), which acted as occluders (so-called ‘phantoms’ [107, p.199]) and colliders to the virtual hand. Correctly placed in front of the camera image plane in Unity this gives the user the impression of actually reaching into the video. For instance, a transparent box was placed in the Unity scene so that it exactly matched the desk in the video image (in size, orientation, and position), occluding the virtual arm when it passed underneath or behind it. This effect can be observed in Fig 4.10 (B), where the hand is cut off exactly at the edge of the desk, and thus appears to be reaching under it. Furthermore, the collider prevents the virtual hand from passing through, giving the user the perception of actually pushing against a solid object. For example, when users lowered their hand so that the virtual hand reached the surface of the desk, it would be hindered at continuing downward even if the users lowered their hand further. This would then lead to a position offset between the real and the virtual hand. However, I observed that users adapted their hand movements and stopped whenever the virtual hand collided. Interestingly none of the participants noticed this fact in the experiment - it simply appeared to confirm what they already expected about the world. The curtain similarly limited the users’ hand motion preventing them from reaching through the fabric or the handle, giving the users the impression of pushing a solid bar left or right.

To further improve the user’s depth perception I implemented virtual shadows that are projected onto the real-world footage. For example, the transparent box that represented the desk in Unity was modified to receive a round shadow from the hand representation, when it was located above it (Fig 4.10, C). A similar box is placed on the virtual floor, allowing the user to better judge the length of the arm before reaching the distant controllable objects.
Figure 4.11: **Comparison of masking techniques.** (A) Inpainting by Teala, (B) Inpainting by Navier-Stokes, (C) *Memory Inpainting*. The masked area is approximately indicated by the light blue outline. Note the tessellated and somewhat distorted appearance of the inpainted area in both the Teala and the Navier-Stokes approach, while in Memory Inpainting the area is visibly smoother, albeit blurred.

**Masking of the real hand**

During pilot studies for the LongArm experiment, participants indicated that they found it difficult to concentrate on the virtual hand, when their own hand was visible at the same time. Consequently I worked on masking the physical hand in the video footage, which is known as *diminished reality* [107, p.227].

To make detection of the hand easier, users were asked to wear a long green stocking on their hand and forearm (see Fig 4.2, C), which allowed detection of the respective region with color-keying using OpenCV\(^\text{10}\).

Three different approaches were compared, by which the pixels of this region were subsequently filled. Firstly, there is the OpenCV inpainting\(^\text{11}\) function, which relies on the Navier-Stokes based method (see Figure 4.11, A). Secondly, I applied a variation of the above approach by Teala [122] (Figure 4.11, B). Both of these approaches involve repetitive sampling of the areas in question, which has an impact on performance. Furthermore these approaches both result in a flickering effect when used on the consecutive frames of a video, since in each frame the areas will be colored in a little differently. Finally I implemented a third approach that I refer to as *Memory Inpainting*, which is better suited for masking regions in video frames (Figure 4.11, C). Here the segmented arm-region is simply filled with old pixel values, which are remembered from the frame before the arm entered the image.

Unfortunately, Memory Inpainting has the disadvantage of causing a visible shift of parts of the image if the camera is in motion or objects move around in the segmented area. The resulting ‘ghosts’ can be seen in Fig 4.12, where the hand holding up two fingers was in view when a green object (which


\(^{11}\)OpenCV inpainting documentation: [http://docs.opencv.org/2.4/modules/photo/doc/inpainting.html](http://docs.opencv.org/2.4/modules/photo/doc/inpainting.html) (last accessed: 28.06.2018)
4.3. IMPLEMENTATION DETAILS

Figure 4.12: **Ghost-effect.** An unwanted effect caused by the *Memory inpainting* method are the ‘ghosts’ of objects that used to be in front of the masked area: even though the hand (A) is moved away, the fingers are ‘remembered’ and remain visible (B). The bounds of the masked area are indicated with the light blue outline. Similarly moving the camera might cause a visible shift of the background texture.

was masked) entered the scene (A). Subsequently, when the hand is moved away, the finger tips remain, since the pixels within the segmented area are not refreshed (B). However, in interaction with the AR system we found that users did not look around a lot while their arm was extended in front of them, and all actuated objects moved sufficiently slowly, so that such ‘pollution’ of remembered pixels was minimal.
Virtual reality system for body ownership

The virtual reality (VR) system described in this chapter served as experimentation platform for the studies presented in the Ownershift [35] (chapter 9) and PhysioMotor [36] (chapter 10) papers. To provide a background for discussing the design and implementation of this system, each of these experiments is shortly described as follows.

Ownershift presents a series of transformations of the virtual hand space (VHS), which reduce the strain of prolonged overhead manipulations in VR that would otherwise require awkward arm positions. Through this experiment we showed that the position and orientation of a virtual hand can be changed quite strongly, while still retaining the sense of body ownership of the virtual arm. This technique allows to better accommodate the user’s overhead task in VR.

PhysioMotor aimed at exploring the contribution of visual-motor synchrony and cardio-visual feedback towards the illusion of owning of a virtual hand in immersive VR. In this user study we evaluated three different virtual hand representations and compared how the type of multisensory stimulation affects our tolerance of unrealistic visual features of the hand. Our findings support that visual-motor synchrony has a strong impact on the body ownership illusion (BOI) and may compensate for lacking connectivity of the virtual

Figure 5.1: Components of the VR system. For the VR system the Display module consists of a VR headset and the Tracking module provides high-fidelity motion tracking of the user’s hand and fingers. Additional modules were added for individual projects, such as the Haptic module and the Physio module, which provided additional feedback to the user.
CHAPTER 5. VR SYSTEM FOR BODY OWNERSHIP

Figure 5.2: Display and Tracking module. (A) HTC Vive headset with LeapMotion sensor (conventional setup), as used in PhysioMotor. (B) HTC Vive headset with double-Leap setup, as used in Ownershift. (C) The double-Leap setup allows hand tracking along line of sight, as well as in front of the user’s body.

5.1 VR system components

The diagram in Fig 5.1 shows the components which form the VR system. The Display module and Tracking module are the central core of the system. Additional modules (Haptic module, Physio module) were added based on specific requirements due to the interaction technique or experimental design. A more detailed description of each component can be found below.

Display module for immersive VR

For the VR system the head mounted display (HMD) was upgraded from the old Oculus Rift DK2 to a headset of the newest generation: An HTC Vive\textsuperscript{1} (Fig 5.2). This provided a wider field of view and higher resolution, at the cost of a slightly heavier device.

Tracking module

To estimate the user’s point of view in the virtual environment, the Vive headset is constantly tracked by two base stations, also known as the Lighthouse tracking system\textsuperscript{2}. Apart from the user’s head, I also wanted to track the user’s hands to animate the virtual hand with natural hand and finger movements. This was supported by the LeapMotion\textsuperscript{3} sensor, which is shown attached to

\textsuperscript{1}Vive: \url{https://www.vive.com} (last accessed: 28.06.2018)
\textsuperscript{2}Vive Lighthouse system: \url{https://www.vive.com/eu/accessory/base-station} (last accessed: 28.06.2018)
\textsuperscript{3}LeapMotion: \url{https://www.leapmotion.com} (last accessed: 28.06.2018)
the headset in Fig 5.2 (A). The installed driver version was Orion (Beta)\(^4\), and there is a Unity plugin that allows easy integration of the technology. Furthermore, LeapMotion provides some useful core assets that include a variety of virtual hands to choose from\(^5\).

The LeapMotion sensor was specifically developed for tracking human hands and allows mid-air interaction with computers through gestures. It emits light in the infrared spectrum and includes two cameras with wide-angle lenses, which capture any light that is reflected from bright surfaces such as our hands. The LeapMotion driver segments this raw footage removing background objects, and converts the remaining bright parts of the image (i.e., any close objects that reflect infrared light similarly as human skin) into 3D data. The hands are then recognized based on some undisclosed algorithms, presumably by matching the 3D data to possible poses of a hand model. This leads to faithful replication of (even subtle) hand movements.

The sensor can be placed on the desk allowing hand tracking from underneath for interaction with desktop applications. Alternately, it can be attached to the front of the HMD (vertically and horizontally centered) providing hand tracking in line of sight for VR applications. Since in the PhysioMotor experiment participants were to experience the virtual hand by looking at it, this was the obvious placement for the sensor (Fig 5.2, A). However, with the sensor attached to the HMD you cannot reach out with your virtual hand and then shortly look away, because by doing so the system will ‘lose’ your hand from view. This proved to be a serious limitation for the Ownershift technique, where we aimed to support interaction with an overhead target, by tracking the hand at waist-level. The problem was that when the users were looking up at the target, the LeapMotion sensor would not be able to track the users’ hands in front of their body.

This limitation was overcome in building a double-Leap setup by mounting a second LeapMotion sensor on the lower-front part of the HMD (Fig 5.2, B). This second sensor was tilted slightly downwards, so that it would provide a view of the space in front of users’ torso, while they were looking straight ahead or upwards. In this way their physical hands could be tracked at waist level, as well as along line of sight (see Fig 5.2, C).

A difficulty I faced when integrating the second LeapMotion was that the driver only supports a single connected sensor. Thus the second device needed to be connected to a separate PC and the data could then be streamed to Unity via the Websockets protocol. This required parsing and converting of the raw tracking data, which was not made easier by the fact that LeapMotion’s Web-

\(^5\)LeapMotion hands module: https://gallery.leapmotion.com/hands-module (last accessed: 28.06.2018)
socket support appears to be deprecated and the documentation is no longer accurate. Once the tracking data was correctly deciphered a heuristic was needed by which the data from both sensors was combined to set the rotations and positions of all the bones in the virtual hand. Since the position of the interactive targets was known and the experiment involved clearly defined interactions, the most simple approach was to switch between sensors based on the position of the interactive panel and the vertical position of the users’ physical right hand. Hence, if the users interacted with either the top or bottom panel with the physical hand in line of sight, tracking data was delivered from the front-facing LeapMotion. However, if the users interacted with the top panel and the vertical hand was shifted up towards the target so that the physical hand was moving at waist level, the system switched to the lower downward-facing sensor.

To compensate for the positional and rotational offset between both sensors, the system was calibrated before use, mapping the coordinate system of the lower sensor to that of the top sensor. This was achieved in two steps: First the position offset between both sensors was evaluated and the origin of the coordinate system of the lower sensor was translated by this vector. Then the rotation offset was evaluated and the coordinate system of the lower sensor was rotated accordingly. To calculate the offsets, tracking data was collected by both sensors while the users moved their hand in circles within the overlapping tracking space. Ten position and rotation samples were collected at regular intervals, which were then averaged. Since the LeapMotion’s tracking space is warped at the edges, the collected tracking data samples were not completely accurate and this calibration process did not cause the virtual hand positions to align perfectly at all times. However, it worked sufficiently well and none of the users remarked on noticing any inconsistencies when switching from one sensor to the other.

In future, to support arbitrary interactions with unknown targets, a more elaborate approach will be needed for tracking the entire reachable space. The LeapMotion tracking data includes a confidence index (according to the documentation a value between 0 and 1 that indicates how well the hand is being tracked), so the intuitive option when using the double-Leap setup would be to always use the sensor with highest confidence. Unfortunately, this value appears to be quite arbitrary and is not updated at all once the hand is lost. For instance, if confidence was 1 (highest confidence) in one frame and the hand leaves the tracking area in the next frame, the value remains 1 even though the hand is no longer tracked. Another obvious solution would be to interpolate or average the position and rotation values delivered by both sensors. But here again a reliable confidence index would be helpful. A further approach would be to somehow fasten the LeapMotion to the user’s forearm, guaranteeing constant tracking of the fingers, while supplying the hand position through an external tracking system such as OptiTrack. Needless to say,
5.1. VR SYSTEM COMPONENTS

Figure 5.3: **Haptic module.** In Ownershift participants received visual and haptic feedback during interaction. (A) Participants could feel a short vibration every time the ball bounced off their finger tip. (B) When touching the panel participants could see a trail of particles following their finger and feel vibrations on their finger tip.

This would lead to a multitude of other challenges, such as how to mount the sensor on the arm without hindering the users’ movements and how to avoid interference between the tracking systems, since both rely on infrared light.

**Haptic module**

To make interactions more convincing in the Ownershift experiment, a vibration motor delivered haptic feedback to the participants’ finger tip in accordance with visual cues. The vibration motor was a Linear Resonant Actuator (LRA\(^6\)), which can provide subtle tactile feedback to the user’s finger tip. This motor was chosen for its small size, lightness, and downward-directed vibration. A Sparkfun Haptic Motor Driver\(^7\) was used to control the LRA through an Arduino Uno\(^8\), which was programmed using the Arduino IDE. During the experiment the LRA was taped to the participant’s finger tip. To keep the device on the hand as light and unobtrusive as possible, the LRA was connected to the Arduino through a long cable (4 meters length) that led up along the user’s arm and then along the cables of the HMD to the Arduino at the PC.

Haptic stimulation occurred in two instances: in the body ownership-elicitation phase and during interaction with the interactive panel. To elicit the illusion of body ownership at the beginning of each trial participants experienced a

---


\(^7\)Sparkfun Haptic Motor Driver: [https://www.sparkfun.com/products/14031](https://www.sparkfun.com/products/14031) (last accessed: 28.06.2018)

\(^8\)Arduino Uno: [https://store.arduino.cc/arduino-uno-rev3](https://store.arduino.cc/arduino-uno-rev3) (last accessed: 28.06.2018)
virtual ball, which appeared to be bouncing up and down on their finger, as shown in Fig 5.3 (A). Each impact of this ball on the virtual fingertip could be felt as a short, subtle vibrations.

When asked about their opinion on the bouncing ball, participants responded that it felt “real” (p2, p6, p9, p15, p16), with one participant explaining that it was the “most ‘real’ experience” during the entire experiment (p2). Other descriptions were “fun” (p2, p4, p9), “cool” (p8, p10, p14), and that they “liked it” (p6, p13). However, some also perceived a slight “delay” or “offset” between visuo-tactile feedback (p3, p12), and that the physics were “not right” (p5, p11, p12). For the questionnaire statement “I had the impression that I could feel the virtual ball on my finger.” 75% of participants (1. quartile) rated 2 or higher (on a scale from -3 “strongly disagree” to 3 “strongly agree”). Ratings for “The virtual ball was convincing.” were nearly as high (median = 2, 1.q = 1). An interesting observation was that, while watching the ball at least 8/16 participants spontaneously started moving their hand up and down in time with the bounces, without being encouraged to do so. When asked about this, they responded that they wanted to make the ball bounce higher (p1, p10, p13, p14) and that it actually “seemed like it went higher” (p7), even though this behavior was not implemented. Some participants also reasoned that it looked or felt “more real” that way (p2, p5).

The placement of the vibration motor on the fingertip further provided the opportunity of delivering haptic feedback for interaction with the interactive panel. During the pursuit tracking task (described in section 5.3) the vibration motor allowed the users to feel whenever the finger tip of the virtual hand touched the interactive panel. Faster movement of the finger across the panel caused stronger vibrations, akin to the haptic sensations when moving your finger over a rough surface. This was accompanied by a touch-visualization consisting of green particles that trailed behind the finger (see Fig 5.3, B), as if seeing a trail of sparkling plankton when reaching into the ocean at night. This visualization was added to provide a more ‘reasonable’ explanation for the haptic feedback, since the panel appeared to have a smooth surface. This should further give the user a clear indication of being in touch with the panel, since judgment of depth was again challenging for most study participants. One participant (p3) reasoned that in real life he would lightly rest his hand on the panel, thus the physical panel would partly support the weight of his hand and guide his stroke. With the virtual panel he needed to pay extra attention to stay on this plane, which is something he was not used to.

In response to what they thought about the interaction with the panel, most participants found the tactile feedback good and helpful (p1, p2, p3, p10, p11) and that it made interaction feel more natural and real (p14, p15). Some mentioned that they felt strongly like they were touching something, even though it did not necessarily feel like a solid panel (p3, p9, p10, p16). For
5.1. VR SYSTEM COMPONENTS

Figure 5.4: **Physio module.** The Physio module involved placement of 3 electrodes on the participants’ torso (A), to record the ECG data using a g.USBamp signal amplifier (B).

instance, p16 said “I didn’t feel like I was touching a surface so much as passing through an electrical field”. Others were less convinced: p6 - “felt like something is vibrating on my finger, rather than I’m touching it”, p7 - “being able to push through panels messes up the entire illusion of there being tactile feedback (of a surface)”, p8 - “it didn’t feel like I was physically interacting”. This is also reflected in the ratings of the questionnaire item “I had the impression that I could feel the panels when I touched them.”, where the median rating was 1 with 75% of all participants responding positively (> 0). Participants indicated even higher agreement to the statement “The tactile feedback when touching the panels matched my movements.” (median = 1.5, I.q = 1). The material of the panel was described as “(a slice of) water” (p3, p5, p15), “foam” (p3, 14), “gel/jello” (p3, p13), “glass” (p11, p15), “a soft LCD display” (p11), and “sand” (p12).

**Physio module**

Part of the multisensory feedback provided in the PhysioMotor experiment consisted of a heartbeat visualization, which was enabled by the **Physio module**.

Three electrodes are attached to the user’s torso (Fig 5.4, A) and Electrocardiogram (ECG) data is collected using a g.USBamp bio-signal amplifier by g.tec\(^9\) (Fig 5.4, B). The signal is then processed in Simulink\(^10\) and either logged for later analysis, or transmitted to Unity for heartbeat detection during run-time.

\(^9\) g.tec: [http://gtec.at](http://gtec.at) (last accessed: 28.06.2018)

CHAPTER 5. VR SYSTEM FOR BODY OWNERSHIP

Cardio-visual feedback

Each heartbeat results in a visible spike in the ECG signal (the QRS complex) [7, 42] and can be detected using a simple threshold. This threshold was calibrated for each participant to guarantee reliable heartbeat detection. The visualization of each heartbeat was implemented as a subtle change of color of the virtual hand [7, 120]: At each detected spike the color of the virtual hand changed to dark red and then slowly faded back to the normal skin tone before the next heartbeat (see Fig 5.5). This resulted in a pulsing effect, as if the blood ejected by the heart was briefly coloring the hand red.

Results of the PhysioMotor experiment show that the cardio-visual feedback was effective, since participants noticed that the hand flashed in the Physio Condition, and were confident that it did not flash in the other conditions. However, participants were not previously informed about the meaning of the color changes and did not recognize the flashing as their own heartbeat. More detailed findings are reported in PhysioMotor [36], attached in chapter 10.

5.2 Elicitation of body ownership in VR

Multisensory stimuli

The Ownershift experiment supported multisensory integration through visuo-proprioceptive, visuo-motor and visuo-tactile stimuli. At the beginning of
5.2. ELICITATION OF BODY OWNERSHIP IN VR

Each trial, there was a body ownership-elicitation phase, in which the real and virtual hands were collocated (the virtual hand space was not yet shifted) and participants were encouraged to move their hand and observe how the virtual hand responded. Furthermore, they experienced a virtual ball bouncing up and down on the tip of their virtual index finger, feeling haptic feedback at every impact (as described in the Haptic module section). During interaction with the panels, users furthermore experienced visuo-motor stimuli from the virtual hand following their movements, as well as visuo-tactile stimuli.

The purpose of the PhysioMotor experiment was to compare different types of multisensory stimuli and their effect on the illusion of owning three virtual hands of varying realism. Therefore three Stimulation Types (conditions) were tested: visuo-motor (Motor Condition), cardio-visual (Physio Condition), and visuo-proprioceptive (Passive Condition) synchrony. In all conditions the virtual hand was nearly collocated with the participants’ real hand. This provided congruent visuo-proprioceptive stimuli and the Passive Condition relied on this alone to elicit the BOI for the virtual hands. The participants were asked to leave their right hand resting on the table top without moving it and no hand or finger tracking was supported. The Physio Condition was similar, however here the participants additionally received cardio-visual feedback, as described in the Physio module section above. Motor Condition supported visuo-motor stimuli through full hand and finger tracking. Participants were instructed to move their right hand and observe how the virtual hand responded. Related studies have found that the illusion of owning an artificial hand could be elicited with each of these combinations of multisensory stimuli. Similarly, we found them all effective for elicitation of ownership of a realistic looking hand. For less realistic hands visuo-motor synchrony showed the most promising results (see chapter 10 for more details).

Virtual user representation

For the Ownershift experiment I aimed to find a hand representation that strongly invoked the body ownership illusion. We adapted a hand model (PepperBaseFull hands) from the LeapMotion Hands Module\textsuperscript{11}, which features a very realistic-looking hand with visible fingernails, creases and lines (hand 301 in Fig 5.6). Our adaptation mainly involved an extension of the arm beyond the elbow, to prevent participants from seeing the end of the arm when reaching out to the interactive panel. This allowed participants to remain under the impression, that the virtual hand was indeed connected to their own body, which earlier work has found critical for the BOI [115, 123]. When explicitly asked for comments on the virtual hand representation in a first pilot study, participants indicated that the hand seemed a little too plump and the fingers

\textsuperscript{11}LeapMotion Hands module: https://gallery.leapmotion.com/hands-module (last accessed: 28.06.2018)
CHAPTER 5. VR SYSTEM FOR BODY OWNERSHIP

were too short. Generally the participants seemed to strongly compare the virtual hand to their own hand and found fault in the mismatch of details such as creases and the size of the fingernails. To explore whether this degree of realism might be within the well known uncanny valley [87], we designed a more abstract virtual hand by reducing the number of polygons and structural details, which resulted in a blocky model of a hand without fine details (Fig 5.6, hand 302). Participants used this abstract hand for interaction in a second pilot study. For comparison they were shortly permitted to try out the more realistic hand at the very end of the experiment. Generally, participants indicated a preference for the more realistic hand. Furthermore, the body ownership ratings were higher in pilot 1 (more realistic hand) than in pilot 2 (abstract hand). Therefore the realistic hand representation was chosen for consecutive experiments. It appears that by asking for feedback on the virtual hand, participants felt encouraged to list differences between the virtual hand and their own, but this did not mean that they felt bothered by these. Perceived similarity also seems to be subject to strong inter-personal differences, since one participant exclaimed that the virtual hand definitely was his own and looked exactly like his own (fingernails and all), and expressed great wonder at how this was possible. This may refer to the phenomenon that the BOI makes the artificial hand seem more similar to our own.

Another important aspect in relation to the hand representation refers to the faithful reproduction of finger movements. Since I permitted the participants to freely move their hand and fingers and observe how the virtual hand responded, they soon found poses which the LeapMotion sensor was not capable of replicating (e.g., when crossing fingers). While the exploration of such limitations seemed to amuse participants, it presumably also weakened their feeling of owning the virtual hand. Therefore I formulated strict instructions for hand motions in the main experiment.

In PhysioMotor we aimed to evaluate how different multisensory stimuli would affect the illusion of owning virtual hands with varying degree of realism. For the experiment we required a hand that was generally rated as very realistic, one that was very unrealistic and one of medium realism. To identify these I conducted an online survey, in which participants were asked to rank five different hands (shown in Fig 5.6) by realism. The three chosen hands (A, B, C) are marked with yellow outlines in Fig 5.6. The most realistic hand representation (hand A) featured the same detailed human hand model with fingernails and creases that was also used in Ownershift. The hand representation of medium realism (hand B) consisted of five long, blocky potrusions that extended from the wrist to the finger tips, similar to the bones in our skeleton. Thus the hand did not actually have a filled-in palm. Some participants described this hand as “carved from wood”. The unrealistic hand representation (hand C) consisted only a thin arm from the elbow to the wrist and claw-like cones at the finger tips. There was no visible connection between
5.3 Implementation details

This section provides details about the Ownershift interaction technique and the tracking task which was used for evaluation in this experiment. Furthermore, it discusses the experimental setup used for PhysioMotor and the threat that was designed to evaluate the virtual hand illusion (VHI).

Implementation of the virtual hand space shift

The Ownershift interaction technique reduces strain during prolonged overhead interaction in VR, by gradually shifting the virtual hand towards the overhead target. While the virtual hand slowly moves upward, the user compensates by moving his own hand further down, reaching a more comfortable position, as shown in Fig 5.7 (A). As described in Fig 5.7 (B), the shift consists of a translation of the VHS along a circular path, which describes the rotation...
CHAPTER 5. VR SYSTEM FOR BODY OWNERSHIP

Figure 5.7: Ownershift technique. (A) Interaction with the top panel begins with 1:1 mapping. Over time the virtual hand is shifted, guiding the users to lower their physical hand. (B) The virtual hand space shift is calculated as a translation along a circular path (based on a rotation around the shoulder by an angle $\alpha$). This leads to an offset $d$ between the real and the virtual hand.

of the user’s hand around the shoulder, by an angle $\alpha$. In other words, the virtual hand ends up where the users’ own hand would be, if they reached up to interact with the overhead target. This leads to a vertical offset $d$ between the physical hand and the virtual hand.

This technique was evaluated in an experiment where an interactive panel was positioned above the participants’ shoulder. This was compared to interaction with the panel in the ‘optimal’ interaction space at waist level [10, 47, 82]. The top and bottom position of the panel was calibrated to the user’s height and the average distance between both positions was $d = 65$ cm. Horizontally the panel was always located slightly to the users’ right, to avoid reaching across their mid-line when interacting with their right hand. In the experiment participants were free to step closer to the panel or move further away, if this felt more comfortable.

The interactive panel consisted of a semi-transparent foreground plane and a dark background plane. This allowed users to reach slightly into the panel, as if it was a “slice of water” (as indicated by participants 3, 5, and 15). The thickness of the panel added a necessary degree of tolerance, since unsupported mid-air interaction makes it difficult to keep the finger exactly on the panel’s surface.
5.3. IMPLEMENTATION DETAILS

Placement of interactive panels

For comparability of overhead interaction with and without shifted VHS in Ownershift it was important that interaction with the panel was similar at both positions, despite varying elevation of the users’ physical hand. Initially the panel was always rotated to face the users’ shoulder, permitting users to interact holding their finger orthogonal to the panel’s surface (see Fig 5.8, A and B). However, when the the panel was at top position and VHS was shifted, the virtual hand appeared to reach up to the top panel while the user’s hand was actually in a similar pose as when interacting with the bottom panel. In other words, while the visual representation of the panel remained overhead, its interactive component was projected down to waist-level (‘projected panel pos.’ in Fig 5.8, C), which made interaction awkward. A solution I pursued, was to rotate the VHS as well as translating it, to make interaction with the hand at waist level identical for both panels. However, this meant forward movements of the real hand would be translated into upward movements of the virtual hand at the top panel. A pilot study soon revealed that this was disorienting, since participants had trouble locating the panel. Next I explored a completely vertical orientation of the panels at both positions. This worked fairly well for interaction with the panel at top position, but was quite awkward when interacting with the panel at the bottom (see Fig 5.8, D-F). This is comparable to the difficulty of writing on a whiteboard at waist level, in contrast to writing above shoulder level - despite the increase in strain, writing on the top part of the whiteboard feels easier. Finally, I supported interaction at both positions using a mixed approach, by orienting the panel at the bottom towards the user’s shoulder and orienting the top one similarly, but slightly more vertical. This made interaction with shifted VHS (Fig 5.8, I) very similar to collocated interaction with the bottom panel (Fig 5.8, G). Unfortunately, for collocated interaction with the panel at the top (Fig 5.8, H) this meant that the upper edge of the panel was somewhat harder to reach, since it was a further away. This also meant that the virtual arm would sometimes partly sink into the semi-transparent panel. However, none of the pilot participants appeared to be bothered by this and this solution seemed preferable to all explored alternatives.

Restricting hand poses for interaction

Moving the hands close to the boundaries of the LeapMotion’s tracking space can lead to judder. Furthermore, unpredictable finger motions can occur for fingers that the sensor cannot see well (e.g., in a perspective when the fingers are occluded by the back of the hand). In the Ownershift experiment this was remedied by ‘locking’ the pose of the virtual hand during interaction with the interactive panel.

In the exemplary pursuit tracking task performed during the user study, par-
Figure 5.8: Orientation of interactive panels. The orientation of panels was varied to find a configuration that supported interaction with the panel at Top and Bottom position with 1:1 mapping, as well as when the hand space was shifted. Both orientation towards the shoulder and vertical orientation led to uncomfortable interaction in some configurations (where the labels are highlighted with orange outlines: C, D and F).
5.3. IMPLEMENTATION DETAILS

Participants were instructed to follow a moving target with their index finger, i.e., they were asked to close their fist and extend their index finger, as if pointing at something. Participants were reminded to remain in this pose and not to move their fingers again until instructed to do so. The fingers of the virtual hand were then ‘frozen’ in this configuration until the tracking task was completed. If participants ignored these instructions, they experienced a non-responsive virtual hand that no longer mimicked their finger movements, which could diminish the illusion of ownership of the virtual hand. However, for the greatest part participants complied and this limitation of movement was not noticed.

For future interaction design, a more flexible approach should be pursued. For instance, the user could choose to lock the virtual hand in self-defined poses.

Implementation of the tracking task

To evaluate interaction performance in the Ownershift experiment, a task was implemented that allowed measurement of task completion time and counting of errors.

The first variant was a Fitts’ law [79] task, consisting of two square targets on a panel. The participants were instructed to tap these targets alternately with their right hand, without missing the targets or accidentally hitting a target twice. They were encouraged to aim for highest possible speed and accuracy. However, a pilot study showed that after a few taps, participants no longer paid much attention to the virtual hand, but simply performed the ballistic movements out of muscle memory. This was problematic since the Ownershift technique, which was to be evaluated, relied on adjustment of the real hand position based on eye-hand coordination.

Hence a second task was designed, which required slower, more controlled movements. In this pursuit tracking task, based on the work by Poulton [98], participants were asked to trace the movements of a small moving target, or marker, on a plane. The position of this marker was computed as the sum of multiple sine waves, resulting in harmonic, yet unpredictable movement. Participants were encouraged to follow the marker as closely as possible with the tip of their index finger.

Performance in the tracking task was evaluated as the ratio of the participant’s tracking error to the error based solely on the movements of the target. In other words, I compared the participants’ tracking error to the error that would have occurred, had the participant’s finger remained in the center of the panel during the entire trial. This was calculated as the root-mean-squared error (RMSE), i.e., the root of the average squared distance between the participant’s index finger and the target over an entire trial.

Results of the Ownershift experiment show that shifting the VHS, as in Own-
**CHAPTER 5. VR SYSTEM FOR BODY OWNERSHIP**

Figure 5.9: **Experimental setup for PhysioMotor.** This study involved minimal experimenter intervention since short questionnaires were answered interactively in VR (A), and participants received pre-recorded instructions through noise-canceling headphones (B). Hereby participants remained fully immersed in VR throughout the entire experiment.

*ershift* condition and *Instant shift* condition, decreased task performance in comparison to conditions with collocation of real and virtual hands (*Bottom* condition, *Top* condition). Interestingly, the participants’ agency ratings indicate that they felt in control despite this increased error. These errors could be addressed by designing targets with at least 3 cm radius. Further details of the analysis are given in the paper, which can be found in chapter 9 [35].

**Experimental setup for minimal external intervention**

When experiencing immersive VR, influences from the physical world, such as the experimenter giving instructions, can reduce immersion and break presence [114]. Since the goal in VHI experiments is to create a convincing impression of the user being part of the virtual environment, it is desirable to keep such disruptive influences to a minimum.

Therefore I designed an experimental setup for PhysioMotor, which allowed to conduct a full experiment with minimal intervention by the experimenter. With this system there was no interaction between participant and experimenter from the point when the participants put on the HMD, until when they took the headset off again. All instructions were prerecorded and delivered to the participant through noise-canceling headphones. At critical points during the experiment, participants could press a button when they had completed a task and wished to continue with the next. In such situations, the system also performed simple checks to ensure that the task was completed correctly and the users’ right hand was where it should be at that time. After each trial, data was collected through an interactive questionnaire. The
5.3. IMPLEMENTATION DETAILS

Figure 5.10: **Threat to the virtual hand.** To evaluate the participant’s reaction to a threat, the virtual hand was pierced by an arrow in the final hand trial. Sound effects of the arrow swooshing through the air and hitting the hand with a thud were added to increase credibility.

Participants saw a question displayed on a virtual panel in front of them and could turn a knob to select their desired response (see Fig 5.9). In this way, participants were constantly engaged in controlling the system and could set the pace by deciding when they wished to continue.

The role of the experimenter was to observe the participant and monitor the system to ensure that no problems occurred. The experimenter intervened only, if the situation commanded it (e.g., if the participants missed part of the instructions the experimenter could cause the instructions to be repeated).

**Threatening the virtual hand**

In the PhysioMotor experiment the participants’ reaction to a threat was recorded as measurement for body ownership. When experiencing body ownership of an artificial hand our physiological responses to a threat towards this hand haven been found to be similar to responses when our real hand is threatened. Hence, measuring the participants’ responses to threat is a reliable method for evaluating the BOI. The treats presented in related studies often consist of stabbing the rubber or virtual hand with a knife [43, 73, 124], or cutting a virtual hand with a circular saw [64].

To avoid showing a floating knife that was cutting the hand, or introducing a new avatar which was wielding the knife, I designed a threat that did not entail the visibility of the agent: an arrow that pierces the virtual hand (shown in
Fig 5.10). Sound effects were added to make the experience more realistic; the arrow could be heard as it was released, whirred through the air and slammed into the table with a thump. I also added a cartoony bleeding animation to highlight the fact that the hand had been ‘hurt’.

The threat was always presented in the last trial at the end of the experiment, since it is assumed that such a threat reaction can only be measured once. As soon as the participant has learned that there is no actual threat, the response would likely diminish. The hand representation shown in this trial was always the most realistic hand, since it was the only hand in the PhysioMotor experiment, which had a connected palm with closed surface that could be pierced. The time of the threat was logged by the system to later evaluate the participants’ heartbeat deceleration in response to this threat.
Research approach

Each of the systems described in chapters 4 and 5 was designed in iterations and evaluated through a series of small pilot studies, followed by a main user study. Due to technical requirements (motion tracking, VR equipment) and to minimize external disturbances during the immersive experience, the experiments and pilot studies discussed in this thesis were controlled lab studies. These took place either at the Department of Computer Science of Aarhus University in Denmark\(^1\) (LongArm and Ownershift), or at the Event Lab\(^2\) at the Department of Clinical Psychology and Psychobiology of the Universität de Barcelona\(^3\) in Spain (PhysioMotor).

The design iterations and evaluation through pilot studies allowed me to test the functionality of the systems, and ensure that the intended interaction was clear to the users. Furthermore, it was important to verify that all instructions and questions of the experiment were understood by the participants, and that the data collection worked as expected. Data collection involved questionnaires in spoken and written form, as well as logging of tracking data, input created by the users, and physiological user data.

Furthermore, pilot studies were a ‘cheap’ way to test our hypotheses. Naturally, the low number of pilot participants (4 per iteration) could only give a vague indication of some effect, but often these findings were confirmed in the subsequent experiment. Results of the experiments provided new insights about the body ownership illusion (BOI) and the effectiveness of the proposed interaction techniques. These findings are valuable for future design of experiences in virtual reality (VR) and augmented reality (AR) that feel more real.

6.1 Experimental procedure for user studies

At the start of each experiment, participants received an explanation of the procedure and were informed about the possible dangers of using VR technology (e.g., motion sickness, epilepsy). They were also informed that they could abandon the study at any time and without giving reasons. Participants then provided written consent, confirming that they participated in the study voluntarily and agreed to let me collect their data in form of questionnaire responses, logged system data, physiological data, as well as audio and video

---
\(^1\)Department of computer science, Aarhus University: [http://cs.au.dk](http://cs.au.dk) (last accessed: 28.06.2018)

\(^2\)Event Lab: [http://event-lab.org](http://event-lab.org) (last accessed: 28.06.2018)

recordings. I then recorded their demographic information (age, gender, handedness, professional background, etc.) and other relevant information, such as their prior experience with computers and VR in particular.

**Study participants**

All studies involved voluntary participants who were recruited mainly among students and staff of the universities at which this research was done. The main recruitment method was by using existing mailing lists and personal contacts, as well as advertisement through posters on campus. All together, the system presented in LongArm was evaluated by 20 participants (5 in pilot, 15 in main study), for the Ownershift experiment I had 29 participants (13 in pilots, 16 in main study) and PhysioMotor involved 76 participants (10 in pilots, 66 in main study).

Efforts were made to ensure diversity among participants (gender, age, professional background). However, limitations in age range and similarities in background could not be avoided due to the typical demographics of the population found on campus. Therefore, experiments performed at Aarhus University mainly involved participants with a background in computer science, whereas the participants recruited in Barcelona were predominantly from the field of psychology. Nevertheless, the level of experience with VR was low for both groups, and all participants that took part in main user studies were naïve to the purpose of the respective experiments.

After completing the experiments, all participants who took part in main user studies received a small compensation for their time. Since financial reimbursement for voluntary participation in user studies is not common practice at the Department of Computer Science in Denmark, the participants were rewarded with sweets and drink vouchers. At the Universitat de Barcelona each participant received a monetary compensation of €10 per hour and an extra €5 bonus if they recruited a friend to participate as well.

**Language of instructions and questionnaires**

Care was taken to ensure that the instructions and questions were clearly understood by all participants. The studies for LongArm and Ownershift were conducted in Denmark, where participants were either native English speakers, or at least fluent in English. Thus verbal and written instructions, as well as questionnaires, were provided in English. Spoken instructions and verbal questionnaires were given by me, and rephrased or repeated if needed.

The research for PhysioMotor was performed mainly in Barcelona, where the two main languages are Castellano and Catalan. Thus, all information sheets, questionnaires and instructions were translated from English, into both local

---

4I am not a native English speaker, but consider myself sufficiently fluent in the language.
6.2. STUDY DESIGN

Figure 6.1: **Independent variables in the LongArm experiment.** (A) arm condition: the realistic virtual hand appears connected to the body through a long virtual arm. (B) hand condition: the realistic virtual hand is disconnected from the body. (C) abstract-hand condition: the virtual hand is unrealistic and disconnected from the body. (D) arm w/o inpainting: same as arm condition, but the user’s real hand remains visible as well.

languages. The instructions in Spanish and Catalan were recorded by a native speaker, while I recorded the instructions in English myself. Each participant could then choose their preferred language for the experiment.

6.2 Study design

**Independent variables**

LongArm involved one within-groups factor: the type of hand representation. By varying the realism of the virtual hand, its visual connectivity to the user’s body, and the simultaneous visibility of the user’s real hand I arrived at four different conditions illustrated in Fig 6.1: arm, hand, abstract-hand and arm w/o inpainting.

In Ownershift I varied the position of the interactive panel (top, bottom) and the shift type (linear, instant, quasi-random, none). With combinations of these factors I arrived at five conditions (Ownershift, Instant shift, Top, Bottom, Control), which are illustrated in Fig 6.2. These conditions were compared in within-groups design.

PhysioMotor had a more complex study design with both a within-groups factor (Visual Appearance) and a between-groups factor (Stimulation Type), as illustrated in Fig 6.3. Hence, participants were divided into three groups, with approximately 20 participants each: one group for each Stimulation Type (visuo-motor, visuo-proprioceptive, or cardio-visual). Each of these groups experienced the same within-group conditions: a realistic hand representation, an abstract hand, and a very unrealistic, disconnected virtual hand.

In LongArm and PhysioMotor the within-groups factors were counterbalanced using Latin square, to minimize effects of order. Latin square means that
Figure 6.2: **Independent variables in the Ownershift experiment.** Five conditions were compared in the Ownershift study. These varied in the position of the panel and the type of shift. Ownershift condition (A) and Instant shift condition (B) involved interaction with the top panel with shifted VHS. Top condition (C) and Bottom condition (D) provided a 1:1 mapping between the user’s real hand and the virtual hand. In Control condition (E) quasi-random VHS shifts were applied to the VHS.

The order of conditions basically stays the same, but conditions are shifted through so that each participant begins at a later condition. A limitation of Latin square, in contrast to fully randomized order, is that there may be carry-over effects, i.e., experiencing one condition may affect the participant’s perception of the next. For instance, if we assume 3 conditions, A, B and C, B follows after A in 2 out of 3 cases (in A-B-C and C-A-B, but not in B-C-A). Hence, if A had a lasting effect that influenced how participants evaluated B, this effect might be visible in the resulting data. This was an issue for the study presented in PhysioMotor (chapter 10), where participants were asked to rate their feeling of body ownership of 3 virtual hands of varying realism. It became evident in early pilot studies that the unrealistic hand was rated more positively if it was the first hand participants saw, compared to when they experienced it last. A reason for this might be the novelty of seeing a virtual hand that they could control in VR, even if it looked unrealistic. Some pilot participants expressed the wish to revise their ratings for the unrealistic hand, once they had something else to compare it to. For this reason the order of conditions was handled differently in PhysioMotor, so that some virtual hands were presented twice: each participant experienced 5 trials, with the hands presented either in the order A-B-C-B-A, or C-B-A-B-C. So the degree of realism was either first increased and then decreased, or vice versa. The first two trials were treated as ‘practice’ trials, which served to introduce the first two hand representations and give participants a frame of reference. The hand representations were then repeated a second time to allow the participants to ‘correct’ their ratings. In the subsequent data analysis I only regarded the last three trials. An alternative to this may be to shortly introduce all conditions
6.2. STUDY DESIGN

Figure 6.3: **Independent variables in the PhysioMotor experiment.** The varied factors were the Stimulation Type (between-groups) and the Visual Appearance (within-groups).

in the beginning, and only then ask participants to rate them. However, this approach may not always be feasible or desirable.

**Dependent variables**

Since the aim of this PhD was the exploration of the BOI in the context of HCI, the degree of perceived body ownership of the virtual hand representation was evaluated in each study. Likewise, the subjective feeling of agency was always evaluated, since it is part of the original BOI questionnaire. Additional factors were evaluated in each experiment.

In the LongArm experiment I additionally evaluated subjective attention (i.e., whether participants focused more on their physical or on the virtual hand during interaction). Furthermore, I asked participants about their preference in regards to the hand representation they used for the given tasks.

In the Ownershift experiment I was interested in the amount of physical strain participants perceived in their arm during interaction with the virtual target. I also calculated task performance and collected subjective self-assessment of the participants’ performance.

The PhysioMotor experiment involved a visualization of the user’s heartbeat on the virtual hand, which made it relevant to evaluate the participants’ degree of...
CHAPTER 6. RESEARCH APPROACH

The methods used to evaluate the dependent variables are described in more detail in section 6.3.

Duration of multisensory stimulation

The duration of multisensory stimulation varies strongly between different BOI experiments. In the original rubber hand illusion (RHI) experiment synchronous visuo-tactile stimulation was delivered for 10 consecutive minutes [15]), after which participants completed the BOI questionnaire. However, this study did not explore how much time it actually took for the BOI to occur. In a later study the duration of visuo-tactile stimulation was less than half as long (4 minutes), and the BOI was still successfully elicited [126]. Other experiments further reduced the duration to approximately 2 minutes [25, 64, 113], and even as little as 90 seconds [62]. Durations for visuo-motor stimulation were similar, for example varying from 3 minutes [73, 105], to just 90 seconds [59]. For eliciting the BOI through vision-only (visuo-proprioceptive congruency), participants spent 1 minute observing a virtual arm from first person perspective [123]. Further, for elicitation of self-identification with and self-location at a distant body representation seen from third person perspective, participants observed a visualization of interoceptive signals (heartbeat) around the virtual body for 6 minutes [7].

The onset of the BOI may depend on the type, number, and quality of multisensory stimuli. For instance, it is imaginable that if the stimuli are very clear and match perfectly, the illusion might start sooner than when we perceive noisy signals, such as a slightly asynchronous stimulation. Furthermore, interpersonal differences may strongly influence the duration of stimulation required to trigger the illusion. For my experiments the duration of multisensory stimulation was adapted based on pilot study results.

The experiment presented in LongArm required the user to interact with the actuated objects in the room for at least 90 seconds with each virtual hand representation. The visuo-motor correlation of the participants’ own movements and the movement of the virtual hand elicited positive ratings for the BOI questionnaire while the virtual hand was visibly connected to the participants’ body through a long virtual arm.

In Ownershift multisensory stimulation was provided for at least 30 seconds in the ownership-elicitation phase (visuo-motor correlation through movements and visuo-tactile stimulation with a virtual ball). During this time participants were explicitly instructed to focus on the virtual hand and reflect on their feelings about it. Thereafter, they spent at least 2 minutes tracing the path of a moving target with the virtual hand, during which they still experienced a certain degree of visuo-motor and visuo-tactile correlation (natural
6.3 Data Collection

PhysioMotor presents an experiment in which participants were instructed to closely focus on each virtual hand for the duration of 1 minute, during which they were asked to reflect on how it felt to see that hand and to imagine, or observe (depending on the condition), how the virtual hand responded when they moved their own hand. Since participants did not need to focus on completing a task during this experiment, the short duration of stimulation seemed adequate.

Despite the varying durations of multisensory stimulation, body ownership ratings are comparable across all three experiments. However, I have not explicitly explored whether the duration of multisensory stimulation has an effect on the BOI, or on the participants' subjective ratings thereof. It seems plausible that the BOI may intensify if the artificial limb is experienced for an extended period of time. Such periods of habituation are common when our bodies change. For example, children move more clumsily when experiencing grow spurts, until they get used to their altered body. Similarly, it can take time to adjust to a new haircut (e.g., lose the habit of unconsciously twirling strands of hair between the fingers), or wearing a party-wig. Further, it stands to reason that mastering the use of a tool can create a strong connection between the tool and our body. If this tool was a virtual hand that strongly deviates from the appearance of our physical hand, we might not feel body ownership of it within 2 minutes, but I argue that prolonged use would increase the likelihood of feeling body ownership of this unrealistic hand. This hypothesis remains to be tested in future work.

6.3 Data collection

This section provides general information about the methods applied to measure the BOI, body awareness, physical strain, and task performance. More information and the full questionnaires can be found in the publications included in Part II.

Evaluation of the body ownership illusion

The BOI is frequently evaluated through subjective indication of how strongly an artificial limb or body is perceived as (part of) the own body. Apart from this, there are several non-subjective measures that provide evidence for the BOI, such as proprioceptive drift and physiological measurements. The following pages provide a description of the evaluation applied in my experiments, as well as a short discussion of additional possibilities for evaluation of the BOI.
CHAPTER 6. RESEARCH APPROACH

Body ownership questionnaire

The most commonly applied method for evaluating the BOI, is based on adaptations of the questionnaire originally conceived by Botvinick and Cohen [15]. Subjects indicate their agreement to each item of this questionnaire on a 5 or 7-point Likert scale (frequently from -3 “strongly disagree” to 3 “strongly agree”), as shown in Fig 6.4.

The LongArm experiment included eight questions that refer to the BOI (Q1-Q4, Q5-Q7, Q9) [34]. Ownershift and PhysioMotor only included four such items (Q1-Q4)[35, 36]. The phrasing of items was adapted for each specific experiment. The full questionnaires are given in each paper respectively.

For the virtual hand illusion (VHI) the main BOI questions refer to whether the virtual hand is perceived as part of the body (e.g., “I felt as if the virtual hand was part of my body.”) and whether it is perceived as the own hand (e.g., “I felt as if the virtual hand was my own hand.”). In PhysioMotor the phrasing of such questions (e.g., Q1) may lead to a different interpretation since it includes the word “might”. For instance a participant could rate it positively thinking “this (realistic) hand looks like it might be part of my body, but it doesn’t feel like it”, which is not the type of meaning that was intended. Therefore it is important to corroborate ratings of such items with others that evaluate the same quality. For example, responses to a negated variant of the same question should be inverse. In PhysioMotor this purpose was served by the item Q2 (“I felt like the virtual hand was NOT part of my body.”).

Multiple other variants of this questionnaire exist, and what they measure depends on how the items are phrased. For instance, most studies adopt the question whether participants had the impression of feeling the touch on the artificial hand from Botvinick’s questionnaire [15], which gives an indication for referral of touch, but not necessarily body ownership. In fact, referral of touch tends to issue stronger responses than the question whether the artificial hand is part of the body [15].

To ensure honest responses, these questionnaires frequently also contain control questions, such as whether participants felt like they owned more than one right hand. This is considered a control question, since generally the BOI has not been found to cause this perception and the expected response is disagreement. In fact, owning an artificial limb usually leads some degree of disowning the own limb [125]. The own body part is thereby ‘replaced’ by a surrogate body part to form a coherent body representation. Presumably this even leads to our body neglecting the affected body part, which has been measured as a drop in temperature of the limb [88]. However, there are cases when this specific item might meet with agreement, for instance when both the real and the artificial hand are visible at the same time. For instance, in the classic RHI it has been shown that we can own a third arm [43]. Similarly,
in the system presented in LongArm participants agreed significantly more to owning two right hands when they could see their own hand at the same time as the virtual hand. Other control questions refer to the appearance of the hand becoming similar to the own or the own hand turning virtual or rubbery.

Agency is frequently evaluated as part of the BOI questionnaire. While there is no clarity as to how these two experiences relate, there is some evidence that prevalence of body ownership makes a agency more likely [17, 136].

It remains to be mentioned, that the use of questionnaires has various limitations. One problem with subjective measures is that the BOI is subject to strong interpersonal differences. More importantly, it has repeatedly been found that some people are non-responders (i.e., they do not experience the BOI or do not admit to experiencing it). Furthermore, some participants tend to be more unsure and respond tentatively, with ratings around the neutral point (i.e., 0 on a scale from -3 to 3). Others have very clear opinions and readily give ratings at the extremes of the scale (-3 and 3). And others again tend respond predominantly positively or more negatively, which may also depend on their mood, their expectations, their experiences, etc. The extents of the Likert scale and the choice of keywords may also affect responses, since they can be perceived differently by individual people. Experimenters should be aware of this and, if possible, should corroborate their findings with objective measures as described in the following sections.

Finally, when analyzing subjective data, interpretation of the responses on the given scales requires some consideration. Since we cannot clearly say that 1 is equally far away from 0 as -1, we must treat the data as ordinal and use non-parametric tests for statistical analysis (e.g., significance tests like Kruskal Wallis or Mann-Whitney test).

Proprioceptive Drift

The illusion of owning a body-external limb is thought to cause a warping of the participant’s proprioceptive perception to include this limb. For instance,
the RHI has been shown to lead to errors when the participants were asked indicate the position of their right hand [15], which is known as proprioceptive drift. Presumably, the stronger the illusion of ownership, the larger the magnitude of proprioceptive drift towards the artificial hand.

Variations of this task include position indication by pointing [15, 54], verbal position indications based on markings on a ruler [126], use of a mouse or controller to move a virtual marker to the perceived position of the hand [36], and pointing towards the position of a distant hand [34, 64].

When evaluating the LongArm system, I measured proprioceptive drift after each trial in a similar fashion as applied by Kilteni et al [64]: participants were asked to extend their arm so that the virtual arm was stretched and the hand came to rest on a location marker in the virtual scene. The screen was faded to black and participants were asked to use their left hand to point to the position of their extended right hand. If the participants were under the illusion of owning the virtual hand, they were expected to point to a position between their real hand and the virtual hand, which was further away. The error was calculated as the difference between the indicated direction (pointing direction of left hand) and a vector from the user’s left hand position to the actual position of the user’s right hand. Hereby, a larger angle means greater drift towards the virtual hand.

In the PhysioMotor experiment the virtual hand was located about 10 cm to the left of the participants real right hand (i.e., the virtual hand was shifted towards the participant’s mid-line), to allow measurement of proprioceptive drift. After each trial the virtual hand was hidden from view and the participants were asked to position a marker where they perceived their real hand to be by turning a knob (see Fig 6.5). To prevent them from remembering the last indicated position (e.g., by the number of revolutions of the knob), this task was repeated twice, once moving the marker from the left side of the table, and once from the right side.

The use of this measurement as an indication of body ownership has been disputed, since it appears that proprioceptive drift can also occur in absence of body ownership. Rohde et al. [104] claim that body ownership and proprioceptive drift must therefore be dissociated and caution against using this measure as a proxy for the BOI. However, they concede that such a reconfiguration of proprioceptive space may be a prerequisite for body ownership (i.e., that when there is body ownership of a displaced hand there is also proprioceptive drift, but not vice versa). Perez et al. [95] further say that proprioceptive drift is not a sensitive measure of ownership, but depends more on the distance between the real and the artificial (virtual) hand.

Unfortunately, the results in the LongArm experiment were inconclusive due to a logging error. The results of PhysioMotor showed no correlation of proprioceptive drift with body ownership ratings. However, across all conditions,
6.3. DATA COLLECTION

Figure 6.5: **Measurement of proprioceptive drift in PhysioMotor**. Participants were asked to positioning a blue marker along an unnumbered ruler to indicate the position of their index finger. The real and virtual hand, which are shown here as yellow and green outlines, were not visible to the participants. Proprioceptive drift was calculated as the error between the indicated and the real position of the participants’ index finger.

Participants exhibited a mean error of 6 cm towards the virtual hand, when indicating the position of their real hand. This may support that drift and body ownership are dissociated. An alternative explanation is that we observed a carry-over effect, which results from the study design: between the trials participants remained immersed in the virtual environment and had no opportunity to recalibrate their proprioceptive perception to the true position of their right hand. This shows that evaluation of the BOI through proprioceptive drift requires careful consideration in the study design, as well as accurately measured and logged data.

**Physiological measurements**

It has been shown that the BOI causes measurable changes in our bodies. For instance, the illusion of owning an artificial hand commonly leads to ‘disowning’ of the own hand, which has been found to cause a reduction of temperature in the affected body part [88]. Ehrsson et al. [29] observed cerebral activity that suggests multisensory integration during BOIs. Furthermore, when under the illusion of owning an artificial limb, we have similar physiological responses to a threat posed to this limb, as when our physical body is threatened. This is measurable in brain activity [31], heart rate [81, 97, 117] and skin conductance response (galvanic skin response) [28, 96, 124, 135]. Further evaluation methods involve observation of increased movement in response to a threat [97, 119], or involuntary muscle activation during observed movement of the virtual arm [116].
CHAPTER 6. RESEARCH APPROACH

Heartbeat deceleration in response to a threat

When we perceive a danger to our body, a deceleration of our heart rate can be detected within approximately 6 seconds after the threat was presented. Prior work has shown that the BOI leads to a similar response, when the artificial body part is threatened [81, 97, 117]. Therefore, in the PhysioMotor experiment I expected to see an effect in the participants’ heart rate shortly after the threat (an arrow piercing the virtual hand, as shown in Fig. 5.10) was presented.

Evaluating body awareness

While it is easy for us to focus on the sound of our own footsteps, or even on our own breathing, it is much harder to perceive body-internal processes that we have no control over (e.g., feeling our heartbeat without taking our pulse). However, listening to our body is something we can learn and practice, for instance through yoga. It has been found that people with strong body awareness (e.g., yogis, dancers, athletes) are less likely to experience the BOI [129]. Inversely, people with a low degree of body awareness have a more malleable body representation, i.e., they will more readily update their internal body representation to incorporate and artificial limb. Furthermore, for people with low body awareness, experiencing the BOI has been found to increase their awareness of body internal physiological processes [37].

Evaluating the degree of body awareness is specially relevant when exploring the effect of visualized interoceptive signals on the BOI, since being able to ‘listen’ to your own heartbeat may help you recognize it when it is visualized on a virtual hand. In this context, I explored two ways of evaluating body awareness: by subjective ratings of the Body Awareness Questionnaire (BAQ) [110], and through a mental tracking method [106, 129], which results in a measure of Interoceptive Sensitivity (IS).

The BAQ includes 18 questions that refer to non-emotional perception of our body. The questions are often divided into four groups referring to detecting changes in our bodily processes, our sleep-wake cycle, predicting bodily reactions and predicting illness [84]. Each statement is rated on a 7-point Likert scale, based on how strongly it applies to the participant. These ratings are then usually summarized into one single factor, by calculating a ratio of averages, resulting in a value between 0 and 1 (or 0% and 100%), where a higher number means a higher degree of body awareness.

IS is calculated based on a heartbeat counting task. Participants are asked to start counting their heartbeats upon a signal, and stop counting at a second signal. During this time they may not take their pulse, but must attempt to feel each beat of their heart in their body and count it. The heartbeats they count are then compared to the number of heartbeats that are measured...
6.3. DATA COLLECTION

with ECG. This results in a score between 0 and 1 that reflects how accurately they counted, i.e., if they counted all heartbeats correctly the score is 1, which means maximal IS. In PhysioMotor the counting task consisted of 1 training trial of 10 seconds, followed by 3 trials of 20, 30 and 55 seconds duration (the order of these trials was counterbalanced). Participants received no indication as to the duration of each trial or how well they did in each. They started counting at the command “start” and at the command “stop” they were asked to indicate the number of heartbeats they counted.

Results show that most participants counted only every second heartbeat (avg. IS = 0.6). Assuming an average heart rate of 100 beats per minute, they would count around 60 heartbeats per minute. This leads to the conjecture that they may have inadvertently been counting seconds, since this is an exercise we are far more familiar with (consider “3 2 1 go” or the countdown to the start of the new year). A further discussion of the results can be found in chapter 10 [36].

Muscle fatigue and strain evaluation

There are many ways to evaluate fatigue during interaction. In connection with manual labor in factories, the RULA score [82] has become popular due to it’s ease of use, since it does not require any complicated measurements nor subjective indications by the workers themselves. This score takes into account the angle at which our limbs are extended from our upper body, and whether our torso is bent over or straight. For instance, repeatedly having to reach up above shoulder level receives a much higher score than arm movements at waist level, reaching across the mid-line is scored higher than when moving the hand at the side of the body, and a posture with the neck tilted up or down receives a higher score than when looking straight ahead. Individual scores for each body part (hands, arms, torso, head, legs) are summed and, based on that, a fatigue score can be looked up in a table.

Unfortunately, the RULA score has discrete steps and does not provide a high resolution measure. This can result in two postures being scored equally or similarly, even though they may have a different perceived level of strain. For instance, in the Ownershift experiment overhead interaction with the top panel would be rated as 4, while interaction with the bottom panel, or with shifted virtual hand would receive a rating of 3. Both of these scores fall into the same category (“low risk”), so this rating does not reflect that prolonged overhead interaction actually became painful, while interaction at waist level felt much more convenient. Montano et al. [86] propose an adjustment of RULA, which allows a more fine-grained evaluation of different poses, by interpolating between ratings based on the angles of individual limbs. Unfortunately, this was not a viable option in my experiment, since the VR system did not provide body tracking data apart from the user’s hand and forearm.
CHAPTER 6. RESEARCH APPROACH

Instead of using RULA, I resorted to a questionnaire that was rated using the BorgCR10 scale [14]. This 16-point scale, ranging from 1 to 10, allows higher resolution for differentiation of lower strain levels, and lower resolution for high strain (see item 5. in Fig 6.4). Strain-related ratings on the BorgCR10 scale have been shown to correlate strongly with muscle activation measures. Based on this questionnaire, participants rated how much strain they experienced in each part of the arm (hand, forearm, upper arm, shoulder) and neck. On the questionnaire, the numbers on of the scale were omitted to avoid confusion, and participants referred only to the textual labels. As expected, the results show significantly higher strain in overhead interaction than for interaction at waist-level. High strain was perceived in particular in the shoulder and upper arm.

A further possibility for evaluation of strain and fatigue are model-based approaches. For example, Hincapié-Ramos et al. propose to measure Consumed Endurance [47], which is the ratio of time we spend in a certain pose, compared to the maximum time that pose could be maintained. By tracking the user’s arm with a Kinect, consumed endurance can be computed during interaction by adding up torques of gravity, the shoulder counteracting gravity and the, arm’s inertia when in motion.

Another example is muscle coactivation clustering by Bachynskyi et al. [10], which relies on a detailed model of 41 muscles in the human torso to examine input regions in respect to muscle strain. The authors identify the ideal interaction space to movements of short to middle length “in the lower right and central part” or reachable space as it has the lowest level of muscle activation resulting in “optimal energy expenditure” and shows improved performance.

The described approaches for objective measure of muscle strain were not integrated in the VR system due to their complexity, and since it would have required tracking of the whole arm or even torso. However, the interactive panels in the Ownershift experiment were positioned based on the classification of different interaction regions, as presented in these papers.

Task performance

In HCI Fitts’ law [79] is commonly used to evaluate interfaces, by measuring how quickly people can navigate through menus or consecutively click different buttons to fulfill a certain task. However, in 3D interaction, such as in immersive VR or AR, Fitts’ law becomes more difficult to apply, since the interfaces we interact with are no longer necessarily based on clicking buttons or navigating list menus. Furthermore, when reaching for objects with our hands in unconstrained 3D space, our hands do not move on linear paths, but rather on arcs around our shoulder sockets. With an interaction that is so distinct from that of moving our mouse cursor around on a flat, rectangular screen, different approaches are needed for evaluating interfaces.
6.3. DATA COLLECTION

One possibility is the use of pursuit tracking tasks, like the one I implemented for the Ownershift experiment. As explained in chapter 5.3, in this task a participant attempts to follow a moving target as closely as possible, either with the cursor, or in our case with the virtual hand. Following the recommendations of Poulton [98], I used the RMSE (root-mean-squared error) to evaluate the users’ task performance.

Apart from this quantitative performance measure, I also collected qualitative data by asking participants whether they felt in control during interaction (agency question in LongArm and Ownershift) and how well they performed the given task (Ownershift). While subjective ratings might not always be reliable and may not always correlate with measured performance, I argue that they are at least equally relevant, since they are part of what defines our experience. Our satisfaction when completing a task is strongly connected to our perceived performance. Doing well makes us feel good, but when we perform badly we are unsatisfied, or even frustrated. The feeling of control over a system is related to our performance, since we can only perform well if we are in control. Furthermore, there is evidence that our perceived feeling of control is in fact a valid indicator for the degree of control we have over a system [133].
Conclusion

This thesis reports on the pursuit of designing virtual body representations that we perceive as our own, and through which we can interact with augmented reality (AR) and virtual reality (VR). By having a virtual body representation the user becomes part of the virtual environment, which dissolves the barrier between real and digital. This renders the computer interface metaphorically invisible.

In my work, I explore body ownership illusions (BOIs) from the perspective of human computer interaction (HCI), thereby combining concepts from the fields of cognitive psychology and interaction design. The described design processes may inform future interaction design, and the prototypes outlined in this work can provide inspiration for future AR and VR systems. These prototypes were evaluated through empirical studies, which are presented in the publications attached in Part II (LongArm [34], Ownershift [35], PhysioMotor [36]) of this thesis.

The following sections present the main conclusions of each of these studies.

7.1 Conclusions of LongArm (chapter 8)

This user study reveals that we can indeed achieve the virtual hand illusion (VHI) in AR. We found the illusion to be robust towards unnatural transformations of the virtual hand, such as stretching the arm to super-natural length [64]. However, in AR it is critical to mask the user’s physical hand, since simultaneous view of the real and virtual hand inhibits the BOI. Further, our results support the conclusion that visual connectivity of the virtual hand to the body is a requirement for the BOI [115, 123]. We can also confirm that the BOI is more likely to occur for hands with a realistic appearance [6, 73]. Last but not least, the presented interaction technique effectively supports direct manipulation of remote actuated objects in the user’s environment, without requiring extensive training.

7.2 Conclusions of Ownershift (chapter 9)

Our findings support that shifting of the virtual hand space (VHS) is indeed a feasible ownership-preserving option for reducing fatigue during mid-air interaction. Compared to related work [60, 75, 92], we found that even stronger position offsets (i.e., vertical offset of 65 cm) between the real and the virtual hand can be tolerated, without breaking the BOI in VR. A gradual increase of the offset may thereby affect the BOI less than an instant shift. The decrease in task performance that may result from such a radical VHS shift could be
addressed by designing targets accordingly (e.g., minimum target size: 3 cm radius).

### 7.3 Conclusions of PhysioMotor (chapter 10)

Body ownership ratings for realistic hands were very high, independent of the type of presented multisensory stimuli. Therefore, our study confirms that visual-proprioceptive stimuli (i.e., vision-only conditions) are sufficient to induce the illusion of ownership for a virtual arm. Similarly, a feeling of agency was reported also in the vision-only conditions: where the virtual hand did not move, but participants were encouraged to imagine what would happen if they moved their hand. This stands in contrast to prior work, which showed that no agency was present without visual-motor synchrony [59]. Furthermore, earlier work suggests that morphological similarity [96, 127, 128] and connectivity [115, 123] are two factors that can inhibit the illusion of ownership for a virtual hand. However, we find that when visual-motor synchrony is given, such incongruities may be tolerated to a higher degree.

### 7.4 Revisiting the research questions

The conclusions given above provide answers for the main research questions, which were presented in the Introduction of this thesis.

**RQ1** Can the illusion of owning a virtual hand be elicited in AR as well as VR?

The VHI can be elicited in AR, under the condition that the virtual hand appears connected to the user’s body and the physical hand is masked. (LongArm)

**RQ2** Can body ownership of a virtual hand be maintained, if this hand is being actively used to manipulate objects?

The VHI illusion persists during interaction in AR and VR (when the virtual hand is no longer the user’s central focus), and is robust to transformations such as increased arm length (LongArm) and spatial shifts between the real and virtual hand (Ownershift).

**RQ3** Can users manipulate their environment with a transformed virtual hand (e.g., unnatural length or position) without extensive training?

A virtual hand that feels like the own allows novices to interact effectively, since users can quickly adjust to transformations of the virtual limb. (LongArm, Ownershift)

**RQ4** To what degree does the type of multisensory stimuli (visuo-motor, visuo-proprioceptive, cardio-visual) impact the body ownership illusion,
7.4. REVISITING THE RESEARCH QUESTIONS

particularly for unrealistic hand representations?

Different multisensory stimuli (visuo-motor, visuo-proprrioceptive, cardio-
visual) are similarly effective in eliciting the BOI of realistic hand rep-
resentations. Visuo-motor synchrony may be able to compensate for
lack of connectivity and thereby achieve stronger BOIs for unrealistic
hands. (PhysioMotor)

Evaluation of the developed prototypes has shown that users can quickly adapt
to interact efficiently with a long virtual arm or a vertically shifted virtual
hand. In the LongArm and Ownershift experiments participants were able to
successfully manipulate their environment within a short time. Study results
further suggest that our mental body representation is more malleable than
expected. When experiencing a virtual arm in AR or VR the BOI is robust
to certain transformations of the virtual hand (e.g., increased arm length,
position offset). This robustness can perhaps be attributed to the integration
of visuo-motor stimuli, which may to some degree even compensate for a lack
of connectivity and realism of a virtual hand. However, it appears that for
realistic-looking virtual hands, vision-only conditions are similarly effective in
eliciting the VHI, as conditions with visuo-motor synchrony.

From related work on BOIs we can summarize:

- synchrony of multisensory stimulation is necessary [15, 22, 59, 66, 115]
- congruent orientation of the artificial limb is important [17, 25, 56, 59, 66,
  115, 127]
- similar positioning of artificial and real limbs is important [60, 75, 92]
- congruent identity of limb is important (also visual resemblance) [43, 66,
  73, 96, 103, 115, 127]
- anatomical plausibility is important (location, connectivity) [66, 80, 103,
  115, 123]
- some mismatches in form and appearance can be tolerated (larger/smaller
  bodies, longer or multiple limbs) [64, 66]

Based on the findings presented above, this list can be extended with the
following points:

- The VHI in AR depends on masking the real arm (LongArm [34])
- Large vertical position offsets between the real and the virtual hand may
  be tolerated, if visuo-motor synchrony is given (Ownershift [35])
- Lack of connectivity may, to some degree, be compensated by visuo-motor
  synchrony (PhysioMotor [36])
CHAPTER 7. CONCLUSION

7.5 Limitations

This thesis provides an exploration of BOIs as a quality of interaction, along with new insights in regards to the robustness of such an illusion towards unrealistic hand representations. Furthermore, it proposes two interaction techniques in which the user’s body is virtually augmented. The systems developed served the purpose of experimentation. However, both the prototypes and the user studies presented have limitations, some of which are discussed here.

Ideally, the AR or VR system should be carefully calibrated to the specific user, taking into account individual characteristics such as inter-ocular distance. Furthermore, to achieve strong embodiment of the virtual body representation the avatar should be adapted to resemble the user as closely as possible; not only in respect to height and overall body size, but in terms of body proportions, gender, skin and hair color, other visible features such as scars, and choice of clothing.

The interaction techniques are further limited, because they are currently restricted to right hand use only. Future work could explore bi-manual interaction, which could facilitate new ways of interacting, or adaptation of the hand transformations by the users themselves. The interaction techniques would benefit from some refinement to provide more immediate control, clearer feedback, and better usability of the systems.

One of the main limitations of the experiments is that they involved artificial tasks. It would be interesting to explore interactions using an owned, virtual hand in more realistic scenarios. Additionally, most test subjects were university students, many of whom were already familiar with either computer science or BOIs. This convenience sample allows for efficiency, but precludes the results from wider generalizability.

7.6 Future work

A fascinating question, which remains to be explored in future work, is whether the BOI changes over time. It would be interesting to examine the effects of (repeated) long-term exposure to a virtual body representation. I hypothesize that, akin to getting used to our body when it changes (e.g., in growth spurts), we would become more familiar with a virtual body over time. This could potentially lead to self-attribution of limbs that are currently deemed too unrealistic to own (e.g., replacing our hand with an appendage that does not morphologically resemble a hand).

Another interesting area for study is to elicit ownership of bodies that are non-human. Could we for instance feel like we owned a bird’s wings or a cat’s clawed paws, if the virtual context provided a plausible scenario? Steptoe et
al. [119] reported ownership of virtual body with a tail, which was seen from third person perspective in a CAVE setup. It would be informative to explore this aspect further in immersive VR or AR, where the virtual body would be experienced from first person perspective.

A further topic to explore, is whether body ownership can be a shared experience. This could be studied by letting two people simultaneously perceive the same virtual body in place of their own. If participants can take turns controlling this body, does the BOI persist even while the other person is in control? This may be useful for rehabilitation purposes, where the therapist and patient ‘share’ a virtual body, or it could ease collaborative control of systems that currently support single users only.

Finally, a perhaps more philosophical question is in how far our perception of inhabiting a place and a body depends on self-location. In VR we perceive ourselves to be where the camera is, so our ‘self’ seems strongly connected to our point of view. Seeing ourselves from third-person perspective has been found to cause out-of-body illusions [28]. What might then happen if we saw the world through a camera that was attached to a drone? When this drone flies away, does our ‘self’ go with it, or do we remain in our physical bodies?

### 7.7 Outlook

I envision a future in which AR and VR technology leaves the lab behind and finds its way into our living rooms and even the outdoors. We might experience virtual content through lightweight glasses, or even contact lenses, while additional sensory cues (e.g., haptics) are provided through wearables or actuators in our environment. User tracking could be provided through wearable devices and smart textiles, or sensors around us.

The advent of VR and AR systems in workplaces and homes will revolutionize the way we interact with computer systems. Instead of pressing buttons, computers might react to our gestures, voice commands, and actions. This evolution of computer systems (from desktop PCs to AR and VR) should lead to a reconsideration of how human-computer interactions are evaluated. Traditionally, good task performance is one of the highest goals of interaction design. But I would urge designers and developers to rethink the criteria by which interaction techniques and systems are judged. I argue that the user’s experience, such as feeling control, presence, and task satisfaction are also important qualities to pursue. For example, I think we should strive to create embodied experiences in AR and VR, providing users with a virtual body representation that they feel to be their own.

---

CHAPTER 7. CONCLUSION

This virtual body representation can then be transformed to augment the human body in a way that is non-invasive; giving us super-human powers such as extending our reach (LongArm) or warping our interaction space (Ownershift). A virtual body representation can be flexibly adapted to specific applications or individual users, an aspect which might also be used to compensate for physical disabilities. For example, used in tandem with robotics, a virtual arm could allow a disabled person to manually open a window or close a door, despite lacking that ability physically. Perhaps an elderly person with limited mobility might regain the ability pour themselves a cup of tea. Transformations of a virtual hand that follows our own hand movements could further provide a revolutionary tool for training, therapy, and rehabilitation [24, 44, 93].

In conclusion, by designing for the BOI we blur the border between the real and the virtual world. Through this innovation we can build computer systems in which the interface becomes imperceptible and we feel like we are, in fact, interacting with the real world.
Part II

Publications
Extending the Body for Interaction with Reality
Tiare Feuchtner and Jörg Müller
Department of Computer Science, Aarhus University, Denmark

Figure 8.1: We present an ownership-preserving direct manipulation technique in augmented reality, which allows interaction with remote devices in a ubiquitous environment with the help of a long virtual arm. While the user’s real hand is close to the body the virtual arm is of normal length (A) and by simply reaching out the user can make it extend to access remote devices in the room. For instance we allow adjusting the height of a table (B), opening and closing a curtain (C) and adjusting the angle of a tilting surface (D).

Abstract
In this paper, we explore how users can control remote devices with a virtual long arm, while preserving the perception that the artificial arm is actually part of their own body. Instead of using pointing, speech, or a remote control, the users’ arm is extended in augmented reality, allowing access to devices that are out of reach. Thus, we allow users to directly manipulate real-world objects from a distance using their bare hands. A core difficulty we focus on is how to maintain ownership for the unnaturally long virtual arm, which is the strong feeling that one’s limbs are actually part of the own body. Fortunately, what the human brain experiences as being part of the own body is very malleable and we find that during interaction the user’s virtual arm can be stretched to more than twice its real length, without breaking the user’s sense of ownership for the virtual limb.

Author Keywords
Ownership; Augmented Reality; Ubiquitous Computing; Virtual Hand Illusion

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

8.1 Introduction

Our rooms and environment are filled with an ever increasing number of interactive devices and embedded computers. These range from permanently installed devices such as automated blinds and smart light bulbs, automatic doors and windows, ventilators and air conditioning, to actuated furniture like height adjusting tables and chairs. How people should best interact with and control this wide range of devices remains one of the big unsolved problems of ubiquitous computing.

The oldest technique is to mount controls on the devices themselves, such as switches, cords, handles etc. Here the main problem is reachability, i.e., the user needs to move around to reach the controls, which are sometimes located in inconvenient locations (e.g., a ceiling-fan). This is improved by mounting controls in some central location, such as light switches arranged at a well accessible position on the wall. However, users still need to access this central location, and it might be difficult to map controls to devices. Another way of solving the reachability problem can be by using remote controls. However, remote controls might not be at hand, and the mapping for multiple devices is once again not trivial.

Recently, more and more devices are controlled by web interfaces or mobile phone apps. Here again the main drawback is that users have to acquire their computer or phone, log in, start the right app, and then select the right device to control. Many other interaction techniques have been proposed, most of which split the problem into \textit{device selection} and \textit{command selection}. Speech interfaces might be useful, but can be socially inappropriate or sensitive to noise and often require the user to learn the command vocabulary. Similarly gestural interfaces require learning of a gesture command language.

We propose going back to the original direct manipulation technique of mounting controls on the devices themselves, and solve the problem of reachability directly. We therefore virtually extend the users’ arms in augmented reality (AR) to enable direct manipulation of the controls or even the devices themselves. Furthermore, to create a better experience we aim at preserving the perception that the arm is actually the users own arm (ownership). Rapid advances in AR and gesture tracking hardware make this technique feasible in the future, where lightweight glasses or even contact lenses could provide the visual virtual overlay, and sensing technology could be ubiquitously embedded in our environment.

In this paper we present an ownership preserving direct manipulation technique with a very long arm and describe the iterative development thereof. We investigate how the fact that the hand is connected to the users’ body, the realism of the hand, and visibility of the user’s real arm, contribute to the perception of ownership in an AR application.
To summarize, the contributions of this paper are (1) a study of the body ownership effect (in particular the moving virtual hand illusion) in AR, (2) an exploration of the limits of body ownership in respect to appearance, realism and connectivity of body parts, and (3) a technique for controlling devices in ubicomp environments that are out of reach.

8.2 Related Work

3D Selection

In order to interact with a ubicomp environment, users need to select and manipulate objects in space. Most of the work in this area has been performed in virtual reality [5, 8, 16]. 3D selection techniques can be categorized by exocentric metaphors (world-in-miniature and automatic scaling) and egocentric metaphors (virtual hand and virtual pointer techniques) [99]. In regards to the latter, virtual pointing is reported to be more effective, more accurate and less strenuous (e.g., through ray casting), while virtual hand can allow a direct transition to the manipulation task (e.g., with a 3D cursor). Our work falls into the category of the virtual hand metaphor. While virtual hands allow to select locations in 3D space, one fundamental problem is that of reach. If control space (the location of the user’s physical hand) and display space (the location of the virtual hand) coincide, objects beyond the physical reach of the user cannot be selected. The key idea to increase reach is to introduce a flexible mapping between control and display, a control to display (C:D) gain.

Control to Display Gain

Poupyrev introduced a C:D gain function that is linear when the hand is close to the body, and increases when the distance of the hand is beyond a certain threshold. In this “Go-Go Interaction Technique” [99] the user controls an abstract hand cursor, which floats in space. The technique increases reach, but not precision. Hindmarsh et al. [39, 48] later applied Go-Go principles in collaborative virtual environments (CVEs), where clicking on distant targets made a humanoid avatar point towards them. The authors found that users had difficulty interpreting other’s pointing gestures and that stretching the pointing arm all the way to the target was helpful. Another class of techniques adjust the C:D gain based on the hand velocity. This is similar to pointer acceleration functions, which follow the same approach for the mapping between mouse and cursor on the desktop [21, 91]. For example, PRISM [40] lowers the C:D gain to enhance precision for object translation and rotation depending on hand speed. Adaptive pointing [70] improves over this approach by simulating absolute pointing behavior. Smoothed pointing [41] presents a further auto-calibrating improvement of the technique.
Distant Reaching in AR and Ubicomp

The need for extending the user’s reach also arises in the context of interaction with interactive tables, large vertical displays, or AR. Some approaches involve using a virtual grabbing tool [1], or laser-style pointing by casting a virtual ray from a touch-pen [94], or the user’s finger [131]. FingARtips [18] is an example for a distant reaching technique in both virtual reality (VR) and augmented reality (AR). This interactive urban planning interface allows users to manipulate objects with a virtual representation of their thumb and index fingers. Two interaction spaces are supported: near space (within arm’s reach), and far space. The manipulation technique for far space includes ray based selection and hand gestures that enable direct manipulation, as discussed in the following section.

Direct Manipulation in AR

While Shneiderman’s principles for direct manipulation [112] originally applied to WIMP interfaces, post-WIMP interfaces that are zoomable, include interactive visualizations instead of icons, and widgets instead of menus, are even more natural and direct [13]. Going another step further AR blurs the line between real and virtual so that direct manipulation may not only apply to virtual objects, but just as well to objects with a physical representation in the real world.

For example the HoloDesk [46], a see-through table surface, allows interaction with 3D objects with the bare hands or everyday objects without the burden of wearing a head mounted display (HMD). It is such a nearly tangible interaction between real and virtual objects, which we strive to achieve. However, instead of making users feel like they can touch something virtual with their real hand, we want to make them think they can touch something real with their virtual hand. This may be possible by changing what people perceive as being part of their body.

To summarize, Pouyrev et al. [99] presents a distant reaching technique in 3D but no evaluation thereof. Hindmarsh et al. [39, 48] later applied this technique in desktop-based CVEs where interaction was supported through mouse input. Many other distant reaching techniques involve tools or rays [1, 94, 131], and while some adopt hand shaped pointers [18], these techniques are all evaluated based on performance (e.g., accuracy, task completion time).

We propose ownership as another measure by which to evaluate interaction, where we focus on experience instead of performance.
8.3 The Psychological Concept of Ownership

Body Representations

We have quite a good idea of what we look like when we walk, we can identify our own shadow or mirror image, and we can navigate past obstacles, e.g., walk through a door without bumping into the door post. For these everyday actions we rely on an internal body representation. Research supports that this representation is malleable and is formed and updated continuously [25]. The body representation is frequently divided into two aspects: the body schema and the body image. The body schema is the knowledge of our own body model and can for instance tell us if an object is within reach, or if we have to duck [80]. We use it to plan our movements and it is updated with each change of posture. We can also update our body schema when using tools so that our ‘peripersonal space’, in which we can act, expands. In contrast, our body image can be described as a mental image of our body, e.g., what it looks like from the outside [25].

Self-attribution: Ownership and Agency

Beyond perceiving the world, we also perceive ourselves and the effects we have on our surroundings. We can therefore differentiate between two types of self-attribution: body ownership and agency. Basically we can define body ownership as the feeling of something being part of our own body, e.g., knowing your arm is yours. Agency on the other hand is the feeling of directly causing changes in the environment. With contradicting findings, the correlation of agency and ownership is still unclear [17, 22, 59, 60, 136].

Body ownership illusions (BOIs)

BOIs [50, 66, 96] involve the self-attribution of external objects to the own body, which can be induced through multisensory integration [20, 80], for instance through concurrent visuo-tactile stimulation of a body part. Such illusions may include ownership of a whole artificial body [81, 96, 117], or just an artificial limb [15, 43, 64, 115]. In our paper, we use the Virtual Hand Illusion (VHI), which is related to the Rubber Hand Illusion (RHI) as first reported on by Botvinick and Cohen in 1998 [15]. The RHI was originally studied using the classic setup of a plastic hand that the user perceives as part of his body. More recently however this ownership effect is frequently explored in a mediated setup, i.e., inducing of the feeling of ownership for a virtual hand [54, 64, 115, 116, 123], and is referred to as VHI. The original RHI study [15] showed that synchronous visuo-tactile stimulation can lead to a sense of ownership of a rubber hand. However, the illusion does not depend on this specific combination of the visual and tactile sense: It can also be elicited through visuo-proprioceptive (or visuo-
motor) integration (without tactile stimulus) [26, 32, 59, 69, 127, 132], or
tactile-proprioceptive integration (without vision) [30]. In our system the
user integrates sensory information from vision and proprioception.
It is important to point out that some incongruities of stimuli have a stronger
disruptive effect on the illusion, than others. To provide a short overview, we
categorize the most inhibiting factors as follows:

- **Matching stimuli**: the stimuli should be synchronous (temporal), collocated
and have the same orientation (e.g., stroking the index finger synchronously
and in the same direction on both the virtual and real hand) [15, 22, 29,
43, 59, 66, 103, 115].

- **Anatomical plausibility**: the artificial limb must be in a plausible posture
(position and orientation) with respect to the body [17, 25, 29, 56, 59, 60,
66, 115, 127].

- **Identity**: the limb must have a familiar appearance and shape [43, 66, 96,
103, 115, 127, 128].

- **Connectivity**: the limb should be visibly connected to the body [115, 123].

If the above criteria are sufficiently fulfilled, the illusion has been shown to be
quite flexible and robust: It was found that the ownership illusion can apply
to supernumerary limbs (e.g., owning a third arm) or unusual limbs [43, 66],
such as arms of different skin color [65], different sizes [11], very long arms
[64], and even tails [119].
It has been shown that, in a VR setup with tactile feedback, ownership for a
passive (non-moving) virtual arm can be preserved, while slowly stretching it
to up to 4 times the normal arm length [64]. It is however unclear if ownership
persists when the arm length is actively controlled by the user and used for
interaction.
Two active VHI studies, where the virtual hand can be controlled and own-
ership is explored during interaction, is the recent work by Lin and Jörg [73]
and Argelaguet et al. [6]. The first presents a VHI comparing 6 different
hand representations of varying realism, which participants used to fend off
spheres flying towards them in VR. The latter compare 3 different hand repre-
sentations, where participants complete simple pick and place operations in a
VR environment. Both confirm that a realistic hand representation positively
impacts ownership.
To give a short summary, most VHI studies are non-interactive, i.e., the vir-
tual hand cannot be controlled by the user. Examples for such are the very
long arm illusion [64], and also the referenced work exploring the effect of
connectivity on ownership [115, 123]. Those VHI studies that do support in-
teraction [6, 73] do not support distant reaching and the realism of the virtual
environment is limited.
We argue that controlling real objects in the environment with our AR setup provides a very different experience. We believe that ownership in AR deserves more attention, since a core difficulty lies in bridging the gap between virtuality and reality in a convincing fashion, which is likely to affect our sense of ownership for a virtual limb.

8.4 Supporting Ownership in Interaction

Designing for Multisensory Integration

To create an illusion of ownership for an artificial or virtual limb, it is necessary that a congruent stimulus is received through multiple senses. In our case we combine vision (seeing the virtual arm move) with proprioception (feeling where your actual arm is). These two stimuli agree in most aspects, such as

Figure 8.2: Optical tracking of the user and the actuated objects in the room is achieved with retro-reflective markers. This picture shows the view of one of the OptiTrack cameras.
CHAPTER 8. LONGARM PUBLICATION

Figure 8.3: (A) Hand in green sleeve, (B) binary mask after chroma keying, (C) *Memory Inpainting.*

how the hand moves and twists and to some degree also what it looks like. It has been shown that the unrealistic effect of the arm becoming longer can, under certain circumstances, be tolerated by the brain [64]. Another factor that could disturb ownership is the delay that occurs between moving the actual hand and seeing the motion of the virtual hand. We have measured this delay to be approximately 100ms which is below the threshold for detecting visuo-motor delays at 150ms [38, 111].

Implementation

Our setup consists of an HMD (Oculus Rift DK2\(^1\)) and we use an OptiTrack\(^2\) system with 10 Flex13 cameras to track the user and the objects in the room. For this purpose retro reflective markers are attached to the user’s head, shoulders, right elbow and hands (the view from one of the cameras can be seen in Figure 8.2). To achieve AR, we converted the Oculus into a see-through HMD by attaching a camera to the front, which provides live video from the user’s point of view. The tracking data and camera stream are both channeled into Unity\(^3\), where we have built a “scene” with virtual representations of the interactive devices in the room and a plane onto which the camera image is projected. Furthermore, this scene contains a virtual full body avatar: a 3D model of a young man that is a standard asset in Unity. Wearing the HMD the user can see the real world captured by the camera, as well as overlaid elements from the virtual world. For instance, the user sees the avatar’s virtual arms as if they were his own. By mapping the movement of the user’s hands to the virtual hands the user receives congruent stimuli from the (virtual) limbs he sees and the (real) limbs he feels (visuo-proprioceptive integration).

\(^1\)https://www3.oculus.com/en-us/dk2/
\(^2\)http://optitrack.com/
\(^3\)https://unity3d.com/
Room layout and interactive devices

To simulate the actuated and sensing environment of the future, an electric curtain and two actuated tables were modified with microcontrollers, allowing them to be controlled from the computer. The layout of the room can be observed in Figure 8.2. The user sits on a stool in the center of the tracking area with a large motorized desk (adjustable in height) located in front of him. The desk is placed out of reach with its front edge 110cm away from the user’s position (i.e., the stool). An electric curtain is mounted on the trusses just behind the table, 220cm away from the user’s position. Slightly to the user’s left and about 150cm away there is a smaller flip-table that can be tilted from horizontal to an angle of about 45°. Markers are placed on the curtain, the desk and the flip-table in order to track their orientation and position. The actuated devices (desk, flip-table and curtain) all have virtual representations in Unity, which are not rendered but simply serve as occluders to the virtual hand when the user reaches underneath or behind them. As a further depth cue we project a shadow from the virtual hand onto these virtual surfaces. Correctly placed in front of the plane with the camera image, this gives the user the impression of actually reaching into the real world, for instance above or underneath the desk, as can be seen in Figure 8.5. Furthermore, we implemented collision detection preventing the hand representation from passing through the virtual representations of real world objects, giving the user the perception of actually pushing against solid objects.

Extending the virtual arm

Our arm-stretch function is based on the Go-Go-Interaction technique [99] by which the user’s virtual arm is stretched non-linearly in relation to the extension of his real arm. While the hand is located close to the body its position is mapped directly to the virtual hand ($R_r$ defines the distance from hand to shoulder). But once the user extends his arm beyond a certain threshold ($R_r > D$), the virtual arm is stretched along a vector drawn from the user’s right shoulder to the right hand, so that the virtual hand is located at a distance

$$R_v = R_r + k \cdot (R_r - D)^p$$

from the user’s body [99]. After some testing, we found $p = 4$ to be a good value, and both $D$ and $k$ are calculated thus to achieve the maximum length of the virtual arm $R_v = 5$. The function is shown in Figure 8.4. That means that when the user’s real arm is fully extended, the virtual hand appears to be 5 meters away from the user’s shoulder.

Apart from the non-linear stretching of the virtual arm, the virtual hand follows the user’s hand motions closely. This allows the user to directly and
CHAPTER 8. LONGARM PUBLICATION

Figure 8.4: The arm-stretch function was implemented based on the Go-Go Interaction Technique. The graph shows the function adjusted to a participant’s arm length, with the exponent \( p = 4 \) and a maximum virtual arm length of \( R_v = 5 \).

naturally interact with objects in the environment that are out of his physical reach.

Three such objects are the above mentioned desk, curtain, and flip-table. Reaching out to the top of the desk with the virtual hand, the user can “push” it down making it move down, and likewise he can raise it again by reaching underneath it, “pushing” it up. The curtain behaves similarly: It has a handle fastened on one side and by “pushing” this handle to the left or right, the user can either open or close the curtain. To adjust the angle of the tilting-table, the user can again just lightly push against the surface of the table to make it rotate in the desired direction.

Masking of the real hand

To enhance ownership, the user’s real right arm is made “invisible” by editing each frame of the video. For this purpose the user wears a green sleeve which covers his hand and lower arm (Figure 8.3, A). Then chroma keying is applied to detect the relevant area in the camera image. After dilating, filtering and
8.5. DESIGN PROCESS

blurring we obtain a binary mask (as shown in Figure 8.3, B), which is then filled with background-colored pixels (Figure 8.3, C). A detailed account of the implementation is made in the following section of this paper.

8.5 Design Process

Our setup and direct manipulation technique were iteratively built and developed involving repeated testing with multiple users and a pilot study (7 participants). During this process, several types of hand representations with varying degree of realism were evaluated. Details about the evaluation methods used in the pilot study can be found in the Evaluation section of this paper, and the most important learnings from this process are described as follows.

The importance of visual depth cues

During early tests it became evident that depth cues were necessary to effectively create the impression of reaching “into” reality with the virtual arm. For this purpose the real world scene was remodeled in Unity to provide realistic collision, occlusion, and shadows. The latter two effects can be observed in Figure 8.5, where the hand is cut off at the edge of the table, and thus appears to be reaching underneath it, or seems to hover above the table projecting a shadow onto it. The addition of these effects was found to reduce the impression of just waving the arm around in front of a flat display, and

Figure 8.5: To provide a better sense of interaction with a 3D environment, we implemented depth cues such as shadows (A) and occlusion (B) of the virtual hand.
CHAPTER 8. LONGARM PUBLICATION

effectively provided a sense of reaching into 3D space instead.

Comparison of inpainting approaches

During the pilot study participants indicated that they found it distracting to be able to see both their real arm and the virtual arm at the same time. They mentioned that this made it more difficult to focus on interacting with the virtual arm. Consequently we removed the real arm by masking the area and replacing it with pixels that are similar to the background color. To fill the masked area three different functions were compared, starting with the OpenCV inpaint\(^4\) function using the Navier-Stokes based method, and the variation thereof by Telea [122]. For performance reasons and to reduce flickering in the masked area we then implemented a third approach, which we find better suited for inpainting video frames. Here the arm-region is simply filled with old pixel values, which are remembered from before the arm entered the image, thus we call it *Memory Inpainting*.

Defining the direction of the arm-stretch

A non-trivial aspect for the implementation of the arm stretch was to determine in which direction the users arm should be extended (i.e., along which vector the virtual hand should be displaced) to provide the best sense of control. Several alternatives were tested with the displacement vector originating either at the user’s head position, the sternum, the right shoulder, or the elbow. Placement at the head would allow pointing in the line of sight, but it also causes the extended arm to move when the user moves his head to look around. Placement at the sternum is more stable in that respect, however it causes the virtual hand to drift closer when the right arm is stretched to the left and further away when the arm is fully extended to the right. Stretching the arm along a vector originating from the elbow results in a very clumsy, long forearm that proved difficult to control. Finally, drawing the vector from the user’s shoulder position makes the hand follow a path that feels fairly natural. We found this to be best suited for our direct manipulation technique.

8.6 Evaluation

We conducted a user study to evaluate our technique and to explore the impact of realism and body-connectivity on the feeling of ownership for the virtual hand representation. Furthermore, we wanted to determine if hiding the real arm is essential for preserving ownership of a virtual arm in AR.

\(^4\)http://docs.opencv.org/2.4/modules/photo/doc/inpainting.html
8.6. EVALUATION

Experimental Design

We evaluated the four conditions *arm* (C1), *hand* (C2), *abstract-hand* (C3), and *arm w/o inpainting* (C4), which we counterbalanced using a Latin Square. The conditions mainly differ in the visual connectivity of the virtual hand to the user’s body, and the hand’s degree of realism and similarity to the user’s actual hand. In C1 the user could see a whole virtual arm and hand replacing his real limb (see Figure 8.6, top left). When the user extended his arm to reach for a distant object, the virtual arm was stretched further leading to a thinning of the limb giving it a rubbery appearance. In C2 the virtual arm was removed and the user could only see the virtual hand floating in the air (Figure 8.6, top right). In C3 this virtual hand was replaced by a hand pointer similar to the one shown in “The Go-Go-Interaction Technique” [99] (Figure 8.6, bottom left). C4 was identical to C1, with the difference that inpainting was deactivated and both the real and virtual arm were simultaneously visible (Figure 8.6, bottom right). Furthermore, all users started with a baseline
in which the virtual arm was of normal length (arm-stretch deactivated), as shown in Figure 8.7.

**Questionnaire**

In each condition the users were asked to answer a questionnaire (see Table 8.1) with statements about ownership (Q1-Q4), agency (Q6,Q7,Q10), and other effects related to the direct manipulation technique. The greatest part of the questionnaire is based on previous work [15, 59, 64], since it has been shown to be a reliable indicator for ownership. We added Q4 to explore if ownership could shift over time and Q8 to find out if seeing both the real and the virtual arm (C4) is distracting. Further, we added Q11-13 to explore in how far ownership involved the expectation of haptic stimuli and Q14 to gather subjective impressions of the experience. All of the statements, except for Q4 and Q14, were to be rated on a 5 point Likert scale (1 “strongly disagree” - 5 “Strongly agree”). After all trials were concluded, the users were asked to indicate their preference of condition. Furthermore, to explore lasting effects of the arm stretch, they were asked if they perceived their right arm to be longer than their left.

**Participants and apparatus**

We performed our user study with a total of 15 participants (4 female), aged 25 to 53 (median 29). Of these, 11 participants claimed to have a computer

---

Figure 8.7: To establish a baseline participants were asked to repeatedly tap 3 virtual cubes with their virtual hand, while their virtual arm was of equal length to their real arm.
8.6. EVALUATION

During the experiment there were times when...

Q1 I felt as if the virtual hand were part of my body.
Q2 I felt as if the virtual hand were my own hand.
Q3 it felt as if I had more than one right hand.
Q4* If you felt like you had more than one right hand, was that simultaneous, or alternating?
Q5 the virtual hand began to look like my real hand.
Q6 when I moved my real hand I expected the virtual hand to move in the same way.
Q7 I adjusted the movement of my real hand according to the movement of the virtual hand.
Q8 During the task I concentrated more on the virtual hand than on my real hand.
Q9 it felt as if my real arm were becoming longer.
Q10 I felt as if I was causing the movement of the desk, table and curtain.
Q11 I was expecting to feel the desk when the virtual hand 'touched' it.
Q12 I was expecting to feel the flip-table when the virtual hand 'touched' it.
Q13 I was expecting to feel the curtain when the virtual hand 'touched' it.
Q14 Describe the interaction in 3 words.

Table 8.1: Participants were asked to respond this questionnaire after each trial (C1-4 and CB). Questions 9 to 14 were omitted in the baseline. All questions, apart from Q4 and Q14, were to be answered by indicating agreement on a 5-point Likert scale (1 strongly disagree, 5 strongly agree). Q4 was only asked if Q3 was rated $\geq 3$.

science-related occupation, and 7 indicated to have experienced AR or VR on more than 3 occasions.

As was described earlier, the apparatus consisted of a see-through HMD that provided a view of the room augmented with virtual content, and motion tracking to allow reaching out with a virtual arm to control remote devices.

**Procedure**

Participants were outfitted with retro-reflective markers before entering the room and were then led into the tracking area with their eyes closed. They were only allowed to open their eyes once they had put on the HMD, so that they could only inspect their surroundings through the head mounted camera.

In the beginning of the study each user was asked to extend his right arm and the virtual avatar was rescaled to match the user’s arm length (measured from the shoulder to the right hand). A large part of the virtual body was made invisible, so that the user could only see the virtual hand representation (depending on the condition).

The participants were then asked to familiarize themselves with their virtual
right arm by observing it move in accordance to their own hand movements. To establish a baseline with a virtual arm of normal length the participants were then instructed to alternately tap three virtual cubes that were displayed in mid-air within comfortable reach of their right hand (see Figure 8.7). The cubes changed color when they were touched and collision detection prevented the hand from passing through, giving the user the impression of solid objects. After 90 seconds of interaction a shortened version of the questionnaire evaluated the user’s sense of ownership for the virtual arm. Thereafter the arm-stretch was activated and the user was shortly introduced to the long virtual arm and how it could be used to control the devices. The participants were allowed to practice interaction with the desk, curtain and flip-table, and then followed four trials with the conditions C1-C4 (Figure 8.6) in counterbalanced order. In every trial the participants were asked to interact with each of the devices, adjusting them as they pleased, resulting in at least 90 seconds of interaction time for each condition. Each trial was again concluded by answering the questionnaire.

8.7 Results

Of the 15 study participants, 3 were excluded from further analysis. One person was disregarded due to technical difficulties during the study, and two were omitted on the basis that they did not give a positive rating (≥ 3) to any of the ownership statements in the baseline. Since ownership is very subjective, great interpersonal differences may exist [25] and some people are less susceptible to the illusion. In previous work participants have been categorized into “responders” and “rejecters” [59], and it is not uncommon to disregard people of the latter group [29]. Since our aim is to explore what effect our different hand representations have on the ownership illusion, we are only interested in the remaining 12 responders. Our main findings are
8.7. RESULTS

Figure 8.9: **Left:** Across all conditions and in the baseline participants maintained strong expectations of the virtual hand following the movement of their real hand. **Center:** Participants indicated that they felt strongly in control of the devices in the environment across all conditions. **Right:** When interacting with the devices the participants adapted the movement of their real hand based on that of the virtual hand. E.g., they stopped when they saw the virtual hand stop at the table’s edge, and adjusted their hand movement to the speed of the table movement.

Evidence of ownership based on questionnaire

The questionnaire responses indicate ownership in the arm condition (C1 with median of agreement: 4 for Q1, 3 for Q2), with similar values as in the baseline (with median of agreement: 4 for both Q1 and Q2) (Figure 8.8, left and center). We performed a Kruskal-Wallis test to compare the conditions and found significantly stronger agreement to Q1 in the baseline, than in the abstract-hand condition (C3) ($p = 0.0084$). No further significance was found for Q1 (C1 vs. C3: $p = 0.0535$, C4 vs. baseline: $p = 0.0845$).

The pattern of stronger ownership indications in C1 than in C2-C4 is also weakly reflected by the ratings for Q9 (felt as if real arm was becoming longer) with median of agreement of 3.5 (Figure 8.8, right). No noticeable ownership occurred for any of the other conditions (C2-C4). It is clear however, that the abstract-hand (C3) was perceived as the most artificial since it got the lowest ratings.

When asked if they felt like they owned more than one hand (Q3) most participants indicated agreement in C4 where both their real arm and the virtual arm were visible (median of agreement: 4, see Figure 8.10, left). This is quite interesting, because in this condition they did not actually indicate ownership based on Q1 (median of agreement: 2), meaning that they felt like they owned a third hand, which was not however part of their body. Here a Kruskal-Wallis test shows a significant difference between the abstract-hand (C3) and the arm w/o inpainting (C4) with $p = 0.0419$. There is no significant difference between C4 and the baseline ($p = 0.0711$), or any of the other conditions.
CHAPTER 8. LONGARM PUBLICATION

Figure 8.10: **Left:** Seeing both the real arm and the virtual arm elicited the feeling of owning more than one right hand. **Center:** During interaction participants concentrated more on the virtual hand than on their real hand, even if it was simultaneously visible (as in C4, arm w/o inpainting). **Right:** After all trials were concluded, we asked each participant to stretch both arms out in front of him and rate the following two statements: P1) *I feel like my right arm is longer than my left.* P2) *I feel like my right arm can reach further than my left.*

**Experiencing an elongated arm**

As mentioned earlier, participants slightly agreed that they felt like their real arm was becoming longer (Q9) in C1 (*arm*) with a median of 3.5, not however in all other conditions (C2-C4). Interestingly however, the interaction throughout all four trials appears to cause a lasting impression of having an elongated arm: After completing all trials we asked the users to stretch out both arms straight in front of them and indicate if they felt that their right arm was longer (P1) and could reach further (P2) than their left arm. Most participants agreed to both of these statements (median of agreement: 4 for both P1 and P2), as shown in the right boxplot of Figure 8.10.

**Agency ratings indicate strong sense of control**

Agency was strong across all conditions with a median of agreement ratings between 4.5 and 5 for both Q6 (*expecting virtual hand to move in the same way as real hand, see Figure 8.9, left*) and Q10 (*feeling of causing movement of devices, see middle boxplot in Figure 8.9*). The participants also indicated that they regulated the movement of their real hand based on the movements of the virtual hand (e.g., adjusting their hand movement to the speed of the moving table that the virtual hand was pressing against), with median of agreement for Q7 between 4 and 5 for all conditions (Figure 8.9, right).
8.7. RESULTS

Strongest focus of attention on virtual hand

In regards to Q8 some participants indicated that their attention switched back and forth between the real and the virtual hand, and in particular in C4 they found it distracting to see the real hand as well and notice the discrepancies. But overall the majority claimed to have focused more strongly on the virtual hand during interaction, with a median of agreement for Q8 between 4 and 5 across all conditions (Figure 8.10, center).

The long arm is the preferred hand representation

After all trials were completed, we asked the participants to compare the long arm (C1, C4) with both the hand (C2) and the abstract-hand (C3) and indicate their preference for each pair. While each hand representation found some supporters, the majority of all users (67%) preferred the arm to the abstract-hand. The reasons include that having the arm made the virtual hand feel more connected to the body, it made navigation easier and was a helpful depth cue. Arguments against it were that it was pretending to be something it was not (the term ‘uncanny valley’ was mentioned). When asked about their preference between arm (C1) and hand (C2) opinions were split with half preferring one and half the other. Arguments for the arm were the same as mentioned earlier and the most common argument against it was that it blocked the user’s view of the target.

Furthermore we asked about their preference concerning seeing their real arm during interaction (arm w/o inpainting, C4), or hiding it through inpainting (C1-C3). Two thirds of all users (67%) preferred not seeing their own arm. Even though our Memory Inpainting creates a blurry area in the image, which still reveals the real hand position, most participants found it less distracting and more immersive than when seeing their real arm at the same time. On the other hand, supporters of having their real arm visible appreciated it as a reference and claimed that it helped them navigate the room with the virtual hand.

Informal description of the interaction

After each trial at the end of the questionnaire, the participants were asked to describe the interaction with three short terms or phrases (Q14). We categorized these descriptions and list the most common in Table 8.2, along with the number of occurrences.

We also asked informally if it was clear to the participants how to interact, if they found it easy, and if the devices reacted as they intended. All of these were confirmed, showing that the participants were able to effectively interact with the system without extensive training.
CHAPTER 8. LONGARM PUBLICATION

Figure 8.11: As a step beyond command languages, direct manipulation interfaces allow users to directly manipulate a representation of data. In the first generation, with Graphical User Interfaces (GUIs), the screen provided the border between the physical and digital worlds. Tangible User Interfaces (TUIs) and Radical Atoms provide a physical representation of the digital world, thereby moving this border into the physical world. Body extension goes beyond these paradigms by moving the border between the physical and the digital \textit{into the user’s own body}.

8.8 Discussion

Study Results

Our user study showed that ownership for a virtual hand in AR can be preserved during interaction with remote devices, if the virtual hand representation is sufficiently realistic and there is a visual connection to the user’s body. We also find that seeing the real hand as well as the virtual hand during interaction disrupts the ownership illusion, thus it is important to hide it from view. In our \textit{arm} condition these requirements were all fulfilled and, in comparison to the other conditions, participants reported highest ownership for this long virtual arm. The long arm was furthermore appreciated because

\begin{table}[h]
\centering
\begin{tabular}{|l|l|}
\hline
\textbf{arm (C1)} & \textbf{hand (C2)} \\
\hline
elastic (5) & realistic (5) \\
weird (3) & disconnected (4) \\
depth cue (3) & hand-like (3) \\
natural (2) & natural (2) \\
easy (2) & powerful (2) \\
\hline
\textbf{abstract-hand (C3)} & \textbf{arm w/o inpainting (C4)} \\
\hline
disconnected (6) & weird (4) \\
mouse cursor (6) & natural (4) \\
unrealistic (5) & easy (4) \\
difficult (5) & reference point (4) \\
powerful (2) & distracting (3) \\
\hline
\end{tabular}
\caption{This table lists the most common categories of descriptions given for each condition. The number of occurrences is indicated next to each tag.}
\end{table}
8.8. DISCUSSION

it helped the participants navigate the environment and judge the depth of their reach. However, users also remarked on it occluding their view.

Applications

While it remains to be evaluated if direct manipulation with an extended body is more efficient than using other techniques, it certainly provides a very different experience. We believe that it is precisely this experience that is the main strength of our technique. Obvious applications include control of the environment for first-time users, and specially also elderly or disabled users who may have less control over their own body than others. This limits their feeling of empowerment and also the perception of their empowerment by others. This empowerment might be increased by the ability to directly manipulate the own environment through a virtual long arm. In the future we envision our technique to be integrated in ordinary glasses and the normal environment of users. Thus, it could be used on a daily basis.

Relationship to Body Implants

Another area that proposes to “extend the body” for interaction is that of implantable electronics [49]. In this field electronic devices are implanted in the human body, e.g., under the skin, to facilitate interaction. Interestingly, while implanted electronics physically become part of the body, they do not necessarily become so from the perspective of the human brain. The brain might well perceive the implanted devices as foreign bodies, if they fail to become part of the body schema. This would presumably happen if the devices fail to provide synchronous sensory feedback to events such as touch and pressure. In contrast, our approach does not manipulate the physical body at all, but merely the brain’s impression of what is part of the body. Whether interfaces are part of the physical body or part of what the brain experiences as the body does not need to correlate.

Body Extension as an Interaction Paradigm

Current interaction paradigms include command languages, agents, and direct manipulation with and without tools. Many current techniques can be subsumed under the “command language” paradigm. Gestural interaction, for example, often consists of a command language of gestures, which are recognized by the computer and execute discrete events. Voice commands and buttons (whether virtual or physical) fall into the same category. The main drawback of commands is that the command vocabulary is limited and needs to be learned by users. Because of the “open mic” problem, the command recognition usually needs to be activated. Target selection, command selection, and command parameterization are usually separate steps.
Virtual (or physical) agents have been proposed to solve the problem of having to learn the command vocabulary. These can range from an invisible voice that can control the room to a mobile robot. Users can converse with these agents in a dialogue. Main drawbacks of agents are that interaction might be cumbersome and socially awkward.

Direct manipulation \cite{112} has been proposed to overcome many shortcomings of command languages, in particular the learning of the command set, thereby producing less cognitive load \cite{53}. Further, direct manipulation interfaces allow users to directly manipulate a representation of data. Instrumental interaction \cite{13} proposes that we can use physical and virtual instruments (tools) to overcome limitations of our own body.

Graphical User Interfaces (GUIs) provided the first instantiation of direct manipulation interfaces, and the screen provided the border between the physical and digital worlds. Tangible User Interfaces (TUIs) and Radical Atoms \cite{55} extend this concept by providing a physical representation of the digital world. Thus, they move the border between physical and digital into the physical world. Body extension goes beyond these paradigms by moving the border between the physical and the digital into the user’s own body (see Figure 8.11).

Combining this extended body with tools is a very interesting direction for future research. Compared to command languages and agents, being able to use the own body with “super powers” might strengthen empowerment, independence, and well-being. This might be particularly beneficial for otherwise less-able users. We believe this might become a very interesting alternative to having to learn commands or interacting with “agents”, e.g., talking to the curtain and asking it to open.

8.9 Conclusion

We have presented a technique where users can directly manipulate devices in their environment through a virtual extension of their real arm. We have shown that the users’ perception of ownership for a long virtual arm can be preserved, if it looks realistic and appears connected to their physical body. Also the visual removal of the users’ real arm aids the illusion. We believe that the virtual extension of the users body is an interesting alternative to gestural or voice commands and virtual agents.
Ownershift: Facilitating Overhead Interaction in Virtual Reality with an Ownership-Preserving Hand Space Shift

Tiare Feuchtner\(^1\) and Jörg Müller\(^2\)

\(^1\)Department of Computer Science, Aarhus University, Denmark
\(^2\)Department of Computer Science, University of Bayreuth, Germany

Figure 9.1: In Ownershift, interaction begins with a 1:1 mapping (A), which allows swiftly reaching towards different virtual targets. If interaction is prolonged, the virtual hand space (VHS) is shifted gradually (B), guiding the user’s real hand into a more comfortable position. Overhead interaction can then continue with reduced strain while retaining similar degrees of task performance and body ownership of the virtual hand (C).

Abstract
We present Ownershift, an interaction technique for easing overhead manipulation in virtual reality, while preserving the illusion that the virtual hand is the user’s own hand. In contrast to previous approaches, this technique does not alter the mapping of the virtual hand position for initial reaching movements towards the target. Instead, the virtual hand space is only shifted gradually if interaction with the overhead target requires an extended amount of time. While users perceive their virtual hand as operating overhead, their physical hand moves gradually to a less strained position at waist level. We evaluated the technique in a user study and show that Ownershift significantly reduces the physical strain of overhead interactions, while only slightly reducing task performance and the sense of body ownership of the virtual hand.

CCS Concepts
- Human-centered computing → Virtual reality; Interaction techniques; Empirical studies in HCI;

9.1 Introduction

We perform most manipulations with our hands. However, the objects that we manipulate are often not positioned optimally, so that it may be uncomfortable or strainful to reach them; in particular if the manipulation takes a fair amount of time. This is especially true for overhead interaction. Fortunately, virtual reality (VR) allows us to shift virtual hands to the overhead target, while our physical (real) hands remain at a lower and much more comfortable position. Such a shift, or translation, of the virtual hand space (VHS) has previously been considered for bringing distant targets into reach. For instance, the Go-Go interaction technique [99] provides a nonlinear transfer function that allows the user to reach distant objects with his virtual hand by simply extending his physical arm. More recently, Erg-O [86] describes a nonlinear mapping by which the interaction space is subdivided into tetrahedrons, within which the position of the virtual hands is dynamically adjusted to make virtual targets more easily accessible. These two techniques demonstrate the considerable potential of such an approach. However, three aspects in particular are not considered in previous work.

First, interaction with overhead targets is very strenuous and could be alleviated by shifting the virtual hand position. Erg-O has the potential to improve interaction with overhead targets by shifting virtual targets downwards (or the virtual hand upwards). However, the maximum offset explored was limited to 10 cm and the authors indicate that participants did not experience any improvement in terms of comfort. We believe that much more radical hand-space transformations are required to facilitate less strainful overhead interaction.

Second, approaches that shift the VHS rely on movement corrections through eye-hand coordination and do not work well for rapid movements. Reaching movements can be separated roughly into an initial ballistic phase and subsequent corrective movements [108, 109]. During the ballistic phase, visual feedback is not processed, and thus a hand space shift would be ineffective. For this reason, both Go-Go and Erg-O depend on interaction with slow, controlled movements. To overcome this challenge, our technique begins with a 1:1 mapping, making it easier to reach swiftly for targets during short interactions. We only shift the VHS after a while, since interactions usually become more physically demanding when continued for a prolonged period of time [108].

Third, one of the most important aspects of interaction in VR, besides task performance, is that the users actually perceive their virtual hands as part of their body. This perception is commonly called the body ownership illusion or Virtual Hand Illusion (VHI) and leads to a more natural interaction, a stronger feeling of plausibility and presence [63], and arguably also a lower cognitive load. The work mentioned above did not explore the feeling of
body ownership regarding the shifted virtual hands. However, in other non-
interactive experiments, the body ownership illusion was reported to break
when the artificial hand was located too far away from the physical hand [60,
75, 92]. It remains for us to explore how the proposed VHS transformations
affect the illusion of “owning” a virtual hand (in the sense of it being part of
one’s body) during interaction in immersive VR.

In this paper, we address the three issues listed above. Accordingly, we present
Ownershift, an interaction technique for reducing the strain of prolonged over-
head tasks in VR. The technique enables users to reach quickly for any target
with a 1:1 mapping, which is beneficial for very fast movements (see Fig 9.1,
A). However, if the overhead reaching pose is maintained for extended peri-
ods of time, the hand space is shifted slowly; the virtual hand moves upward
slowly, gently guiding the user to gradually lower his own hand in order to
maintain alignment with the target (Fig 9.1, B). This leads to a less physi-
cally strained pose for prolonged interaction (Fig. 9.1, C). With our work, we
contribute (1) an interaction technique that effectively reduces fatigue during
overhead interaction in VR, (2) an evaluation of this technique in a user study
comparing it with an instant shift and unaided overhead interaction, (3) ev-
dence of the body ownership illusion, despite a large vertical offset between
the real and the virtual hand.

The following are the main lessons learned from our user study: 1. The
proposed VHS shift successfully reduces strain during overhead interaction,
while task performance and ownership of the virtual hand are only affected
slightly. 2. Ownershift is preferable to an instant shift, since, with similar
task performance, the gradual shift allows initial ballistic reaching towards the
target, is less conspicuous, and perceived as less disorienting and disruptive.
3. The virtual hand illusion seems to be more robust to vertical position offsets
between the real and virtual hand than was previously believed.

9.2 Related Work

Reducing fatigue during mid-air interaction

Interaction with vertical large screen displays, as well as mid-air interaction,
have been found to be quite fatiguing [9, 10], colloquially termed the “Gorilla
Arm”. With the goal of identifying the optimal regions for interaction and de-
signing low-fatigue interfaces, extensive work on dynamic models has explored
this topic. For instance, RULA (rapid upper limb assessment) [82] provides
a system for scoring the muscular effort associated with various different pos-
tures, based on the orientation of the arms and upper body. According to this
method, the least strenuous working posture is one with the hands in front
of the body at waist level, the elbows slightly bent, upper arms relaxed and
torso straight. Similarly, Hincapié-Ramos et al. [47] identified interactions
as least strenuous when hand positions allow bent elbows and are mid-way between the waist and shoulder. Bachynskyi et al. [10] recommended short or medium-length movements in the near and lower part of reachable space.

Hand space transformations

The Go-Go interaction technique [99] is one of the first examples of transformations of the hand space, where users could move the virtual hand beyond their real hand reach so as to access remote virtual objects in immersive VR. The virtual space can also be shifted without the user noticing, for instance, when virtually extending the interaction area by Redirected Walking [67, 101], or using a single haptic proxy for multiple virtual objects through Haptic Retargeting [8, 23]. In this context, Kohli et al. [68] found that performance in a multi-directional tapping task on a haptic proxy was comparable, with and without warping of the virtual space (i.e., same vs. different orientation of the physical and virtual surfaces). Furthermore, Burns et al. [19] indicate that an offset between the real and virtual hand may be readily accepted in order to preserve a believable visualization of the virtual hand, with respect to other virtual objects (e.g., no interpenetration of the virtual hand with other objects). Hence, warping the virtual world and body can enable us to avoid interpenetration of the virtual hand with other virtual objects, or effectively make virtual controls more easily accessible and interaction less tiring, as exemplified by Erg-O [86].

With the aim of making retargeting unnoticeable, most techniques discussed above apply only small shifts (e.g., max. 10 cm [86]). In exploring larger offsets, it has been found that deviations of up to $40^\circ$ were “bearable” for retargeting [23]. In another experiment, horizontal offsets of up to 76 cm were applied when reaching for a virtual target and its haptic proxy [45]. These researchers found that instantly applying the shift was less disorienting and led to better performance than when interpolating the offset while reaching (i.e., retargeting).

In this paper, we explore the effect of applying a large vertical shift (avg. 65 cm) to the VHS, to ease overhead interaction. In contrast, most related systems apply small horizontal shifts for the purpose of guiding the user’s hand to haptic props. Additionally, all approaches discussed above take advantage of the dominance of vision over proprioception, and our natural reliance on visual cues when navigating towards a target (eye-hand coordination). Unfortunately, this does not work well during initial ballistic hand movements towards a target. With our technique, we address this challenge by beginning with a 1:1 mapping and only gradually applying a shift for prolonged interaction. We are not aware of any previous work exploring shifts that are applied gradually over time (vs. interpolation based on the real hand position [8, 23, 86, 99]), and the implications this may have for interaction in VR.
Body ownership illusions

We experience the world with our senses and interpret what we perceive through multisensory integration [20, 80]. However, this sometimes leads to misinterpretation, such as the ventriloquist effect [4]. Another such effect is the body ownership illusion. Body ownership refers to the conviction that your hand, for example, is part of your body, and the illusion of body ownership involves self-attribution of a body-external object. This is probably best known through the rubber hand illusion (RHI) [15]. To elicit this illusion, a rubber hand and a person’s hidden real hand are stroked synchronously with a brush. Since the touches felt can only be observed on the rubber hand, this leads to a referral of touch and the interpretation that the rubber hand must be part of the own body. There appears to be some interplay between body ownership and agency [17, 22, 59, 60, 136]. However, it is important to be aware of the distinction between these two terms. Agency, or the sense of control, refers to a feeling of causing a change in the world. For instance, we feel agency when we click something with our mouse cursor, but we will not feel body ownership of the cursor.

Several variations of the RHI experiment have been performed, in order to explore the processes behind this illusion [29, 32]. While the classic RHI experiment relied on synchronous visuo-tactile stimuli, it has been found that the illusion can be achieved through multisensory integration of multiple different senses, such as tactile-proprioceptive [30], visuo-proprioceptive, and visuo-motor [26, 32, 59, 69, 127, 132]. The latter refers to an artificial hand moving in correspondence to the real hand and has been termed the Moving RHI. The RHI has been replicated in VR, as the illusion of owning a virtual hand [115], which is frequently referred to as the virtual hand illusion (VHI) [6, 73].

There is a large body of work exploring the limitations of body ownership illusions: from congruency of stimuli [17, 22, 25, 29, 111, 127], to mismatches in the appearance, morphology, and connectivity of the virtual limb [96, 103, 115, 123, 128]. It has also been found that the body ownership illusion is affected by the distance between the real and artificial hand, with Kalckert and Ehrsson reporting no body ownership beyond a vertical displacement of 27.5 cm [60]. Lloyd [75] and Nierula et al. [92] confirm, that body ownership is similarly weakened by horizontal displacement. However, while the authors do not discuss absolute ratings, it is clearly evident from their figures, that they observed positive ratings for referral of touch at 67.5 cm [75], as well as some evidence of the VHI at 30 cm displacement [92]. Furthermore, the body ownership illusion proved surprisingly robust to displacements away from the user, as explored for connected virtual arms, that were extended up to 3 times their normal length [34, 64]. The latter examples, which provide evidence of ownership, despite displacement of the virtual hand, have in common that they
CHAPTER 9. OWNERSHIFT PUBLICATION

were realized in immersive virtual environments. This supports our hypothesis that users may be more tolerant of discrepancies between the real and the virtual hand in immersive VR, due to the lack of visual cues from the real world and their real body. Hence, this paper takes a step towards exploring how robust the VHI is towards large vertical position offsets and extends the analysis of how this could be leveraged to improve interactions.

Figure 9.2: A VHS shift results in a vertical position offset $d$ between the real and virtual hand. We arrive at this offset by translating the virtual hands on a circular path, based on a rotation ($\alpha$) around the user’s shoulder on the sagittal plane. To achieve a more realistic posture, the virtual forearm is always oriented towards the user’s elbow.

Figure 9.3: (A) Two Leap Motion sensors were attached to the HMD to support hand tracking, both in the line of sight and in front of the user’s body. (B) The resulting tracking volumes are indicated with lines in light green (front-facing sensor) and blue (downward-facing sensor). (C) Haptic feedback was provided through a small vibration motor that was taped to the index finger. (D) A classic pursuit tracking task was used to evaluate task performance; participants followed a target (light circle) of 1 cm radius, which moved across a 30x30 cm panel with quasi-random motion.
9.3 Ownershift

Transformations of the virtual hand space

We define the virtual hand space (VHS) as the interaction space with its origin at the center of the virtual hand representation, which can be transformed with respect to the user. It could for instance be scaled, rotated, or shifted. A shift of the VHS results in a location mismatch between the participant’s real and virtual hand, as depicted in Fig. 9.2. This allows users to access difficult-to-reach targets with their virtual hand representation, while keeping their own hand in a comfortable position. Apart from this shift, the tracking of the user’s hand remains unmodified, allowing accurate movement of the virtual hand and fingers.

Our work aims to explore whether transformations can be applied to the virtual hand space, so as to render interaction less tiring, while preserving the illusion of ownership for the virtual hand. The shifts we explore are limited elevation changes of the virtual hand. Our focus on vertical shifts is due to the assumption that when interacting with a target off to one side, we can reorient our torso, or even our whole body, towards it to make interaction more comfortable. For distant objects, we are usually able to walk towards them. However, when manipulating an object above shoulder level, there is no option of repositioning ourselves to make this less strenuous.

Facilitating overhead interaction

We explore a VHS shift to allow the virtual hand to be aligned with an overhead target, while the user holds his physical hand at waist level. This vertical position-offset of magnitude \( d \) is the result of a translation of the VHS along a circular path, based on a rotation \( \alpha \) around the user’s right shoulder (Fig. 9.2). Such an arm-rotation in the sagittal plane around the mediolateral axis (i.e., the axis traversing both shoulders), corresponds to shoulder flexion. Note that this technique ensures a consistent mapping of the hand movements, albeit with the virtual hand translated to a different position.

We propose this shift to be applied gradually, and only during prolonged interaction with an overhead target. At the beginning, the real and virtual hands are collocated (1:1 mapping), allowing the user to reach rapidly for the target. This is important, because when we reach for a target with quick movements, we rely less on eye-hand coordination and would not be capable of correcting for a large shift. However, when interaction with the same target is continued involving smaller hand movements, the VHS can be transformed gradually. While this shift is applied, the user must continuously adjust the position of his real hand, in order to keep the virtual hand aligned with the target, so that his physical hand is guided automatically to a more comfortable posture. If interaction with the target requires a certain degree of hand motion, the
shift of the VHS is masked and becomes almost unnoticeable. To reset the shift, our current solution entails detecting when the hand leaves the user’s field of view, upon which the system switches back instantly to a 1:1 mapping. This allows the user to reach for other targets in the interaction space, and if interaction is prolonged, a new VHS shift is applied gradually.

![Figure 9.4: The five conditions varied in panel position, and the type of VHS shift. From left to right: Ownershift condition (O) with gradually applied shift and panel located at the top; Instant shift condition (I) with instantly applied shift and top panel; Top condition (T) with collocated hands and panel at the top; Bottom condition (B) with collocated hands and panel at the bottom; Control condition (C) with quasi-random shifts and bottom panel. The virtual hand is overlaid and highlighted with a yellow outline.](image)

9.4 System design

To explore the effects of shifting the VHS for overhead interaction, we implemented a prototype in Unity3D\(^1\). We use a state-of-the-art HMD (HTC Vive\(^2\)) to provide an immersive virtual environment. Furthermore, the system supports full hand tracking and haptic feedback.

Supporting hand tracking in reachable space

For convincing hand and finger tracking, we rely on the Leap Motion\(^3\) sensor. In our design process, we discovered that with the default configuration of using a single Leap Motion sensor attached to the HMD, we could track the hand well along the line of sight. However, this would not support interaction with a shifted VHS, since the participant would then be looking slightly upwards at an overhead target, while moving his physical hand at waist level. Thus, we decided to increase the tracking area with a second sensor. The

---

\(^1\)https://unity3d.com

\(^2\)https://www.vive.com

\(^3\)https://www.leapmotion.com
most stable solution for this proved to be mounting both sensors on the front of the HMD, one centered and the other at the lower edge, as shown in Fig 9.3 (A). This allows us to additionally track the interaction space in front of the user’s body, while looking at an overhead target (Fig 9.3, B).

Since hand tracking with the LeapMotion sensor relies on the emission of infrared light, multiple, simultaneously active sensors cause interference, which causes the virtual hand to shake. Therefore, it is preferable to have only one sensor active at a time. Hence, the system switched from one sensor to another, when the VHS was shifted. This was based simply on the known location of the interactive target and the user’s hand (i.e., when the user’s hand was about to leave the tracking area of the front-facing sensor due to the VHS shift, the downward-facing sensor was activated, and as soon as it registered the user’s hand, the first was then deactivated). The coordinate systems of both sensors were mapped on top of each other in an initial calibration step, which made the transition smooth and unnoticeable to the user. It remains for future work to devise a more flexible approach for switching dynamically between multiple LeapMotion sensors.

Design of visual and haptic feedback

To make interactions more convincing, visual and tactile feedback is provided whenever the participant’s index finger touches a virtual object, e.g., a bouncing ball (Fig 9.3, C), or an interactive panel (Fig 9.3, D). A trail of particles follows the finger’s path on the object’s surface and the intensity of the tactile stimulus on the fingertip varies continuously with the movement speed. Aiming to mimic the feeling of moving your finger lightly over an uneven surface, it vibrates more strongly when moving faster and ceases to vibrate when the finger rests (motionlessly) on the surface. The tactile feedback is provided through a small vibration motor (LRA\(^4\)) that is attached to the tip of the user’s index finger. The LRA is controlled through an Arduino Uno\(^5\) and Sparkfun haptic motor driver\(^6\), which receives commands from Unity via serial communication.

9.5 Experiment

To evaluate the Ownershift interaction technique, we conducted an experiment comparing the linear shift to an instant VHS shift, and to unaided overhead interaction (no shift: collocation of real and virtual hands).

\(^4\)https://www.precisionmicrodrives.com/vibration-motors/
linear-resonant-actuators-lras
\(^5\)https://store.arduino.cc/arduino-uno-rev3
\(^6\)https://www.sparkfun.com/products/14031
CHAPTER 9. OWNERSHIFT PUBLICATION

Study Design

We performed a user study to evaluate our interaction technique with respect to the following dependent variables: physical strain, the illusion of ownership, and task performance. We designed five conditions varying the independent variables: type of shift and panel position (see Table 9.1).

<table>
<thead>
<tr>
<th>Shift</th>
<th>linear</th>
<th>instant</th>
<th>none</th>
<th>quasi-random</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel top</td>
<td>O</td>
<td>I</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>Panel bottom</td>
<td>-</td>
<td>-</td>
<td>B</td>
<td>C</td>
</tr>
</tbody>
</table>

Table 9.1: The study design consisted of five conditions: Ownershift (O), Instant shift (I), Top (T), Bottom (B), and Control (C). These varied in panel position and the shift type (independent variables).

The conditions are depicted in Fig 9.4. In the Ownershift condition (O), the user started out reaching overhead to track the target on the top panel (1:1 mapping; collocation of real and virtual hand). The shift was then increased linearly over a duration of 45 seconds, slowly easing the user into a more comfortable position with his hand at waist level. This duration resulted from pilot studies, since this shift speed was found to be barely noticeable and yet allowed the full offset to be reached in less than 1 minute (ensuring 1 minute of interaction with full offset in the 2-minute tracking task explained below). The gradual shift was compared to the Instant shift condition (I), in which the same VHS shift was applied, albeit instantly. This required the user to first locate the virtual hand at its elevated position and adapt to the offset immediately, before being able to interact with the top panel. An instant shift was chosen over an interpolation-based approach (e.g., Go-Go [99], or Haptic Retargeting [8]), since this has been shown to lead to better performance in reaching-tasks with large offsets [45]. Both approaches with shifted VHS were explored in contrast to the Top condition (T), which required unaided overhead interaction with collocated hands (no VHS shift). Furthermore, to collect a baseline, we designed a Bottom condition (B) with 1:1 mapping, for which the panel is easily accessible at waist-level, and an asynchronous variant thereof, the Control condition (C). In the latter, quasi-random shifts were applied to the VHS, with the purpose of disrupting the sense of body ownership and control, and leading to sub-optimal task performance. These quasi-random shifts are explained in more detail in the respective section below.

We conducted the experiment with respect to the following hypotheses: Compared to a collocation of the virtual and real hand (Top condition), shifting the VHS (Ownershift and Instant shift condition) reduces strain in an overhead tracking task (H1). The illusion of owning the virtual hand can be maintained,
9.5. EXPERIMENT

despite the large shift of the VHS (avg. 65 cm) in Ownershift condition (H2). Shifting the VHS (Instant shift and Ownershift condition) leads to similar task performance in an overhead tracking task, compared to no shift (Top condition) (H3). Based on these hypotheses, we aimed to find out whether it was beneficial to apply a VHS shift gradually, instead of instantly (RQ1). Another research question was whether task performance deteriorated while the linear shift was applied, since the user additionally needed to compensate for the adjustment of the VHS while following the target (RQ2).

Positioning of interactive panels

The interactive panel measured 30x30 cm and was positioned within comfortable reach in front of the user. Horizontally, the panel was aligned with the user’s right shoulder to prevent participants from reaching across their mid-line during interaction. For the panel’s vertical placement, we defined a bottom position and a top position. In bottom position, the panel was easily accessible at waist level, while in top position, the panel was centered approximately at eye level. For the latter, the user needed to reach overhead when required to access the panel’s upper edge. Panel positions were dynamically adjusted to the user’s height (bottom position avg.: 112 cm, top position avg.: 178 cm).

Design of the pursuit tracking task

To evaluate task performance, we chose a classical pursuit tracking task in 2D, with quasi-random target motion. An overview of tracking tasks for interface evaluation is provided by Poulton [98]. In the implemented task, a round target of 1 cm radius moved on the 30x30 cm panel (see Fig 9.3, D). Quasi-random motion in x and y direction was generated by a sum of 4 sinusoids. Angular frequencies and phase shifts, which were adapted over the course of multiple pilot studies, are provided in Table 9.1 in the appendix. The user’s objective was to keep the tip of his index finger centered on the target at all times.

Quasi-random VHS shift

The quasi-random offsets applied to the virtual hand in the Control condition were calculated similarly as a sum of sinusoids. This was intended to give participants the impression that the virtual hand was uncontrollable. The phase shifts and angular frequencies, in Table 9.2 of the appendix, were chosen on the basis of a pilot study. In trying to compensate for the movements of the virtual hand, the participants were engaging in a kind of compensatory tracking task, in addition to the pursuit tracking task (following the dot). This has been shown to significantly reduce task performance [98].
CHAPTER 9. OWNERSHIFT PUBLICATION

Data collection

Data was collected both by logging tracking data for task performance (i.e., the participant’s ability to follow the moving target) and through subjective questionnaire ratings for body ownership, agency, and physical strain.

Tracking task error

During the trial, the location of the target and the user’s fingertip were recorded in each frame. The position of the index finger was projected onto the panel’s surface, so as to cope with the challenge of depth perception and mid-air interaction, which occasionally led participants to lose touch with the panel. The task error was then calculated as the root-mean-squared error ($\text{rmse}$) [98].

In order to improve the comparability of our results with related work, we used relative errors, as recommended by Poulton [98]. In other words, we analyzed participant performance by comparing the task error ($\text{rmse}$) with the error that would occur if the participant had held his finger in the center of the panel during the entire trial ($refE$).

This results in an error ratio, or percentage, with lower values describing a smaller error and therefore better task performance. For more details, refer to equations (9.1), (9.2) and (9.3) in the appendix.

Questionnaire

After each trial, participants were asked to respond to a questionnaire with nine items (see Table 9.3 in the appendix). The first four items of the questionnaire (Q1-Q4) aimed at evaluating the participant’s feeling of body ownership and agency of the virtual hand. This type of questionnaire is very commonly used to evaluate body ownership, and the items mentioned are based on the original ownership questionnaire [15], as well as later adaptations [59, 64, 115]. Ratings were given on a 7-point Likert scale with values ranging from -3 (‘strongly disagree’) to 3 (‘strongly agree’).

The next five questions (Q5-Q9) prompted the participant to indicate the amount of physical strain they felt during the tracking task in their right shoulder, upper arm, forearm, hand and neck. These body parts were chosen, because they were the most affected by the task, since the different conditions influenced neck-tilt and the degree to which the arm needed to be raised. Strain intensity was rated on the Borg-CR10 scale [14], which has been found to correlate strongly with other endurance metrics, such as EMG [47]. Furthermore, we were interested mainly in the participants’ perceptions of strain (rather than an objective measure), since this affects their overall experience.
9.5. EXPERIMENT

Participants
We recruited 16 right-handed volunteers (4 female) to participate in the experiment. Participants were between 21 and 38 years of age (mean = 28.5, sd = 5.09). All had normal or corrected-to-normal vision and did not suffer from any shoulder injury, or pain when lifting their right arm. Most participants were students or staff at the department and 12 worked in computer science related fields. Nevertheless, previous VR experience was low (median = 2, sd = 1.05), which translates to “1-3 times” on a scale from 1 (“never”) to 5 (“10 or more times”). All participants gave informed consent and were compensated for their time.

Procedure
At the beginning of the experiment, a small vibration motor was attached to the tip of the index finger of the participants’ right hand and the experimenter helped them put on the HMD. The participants found themselves in a bright, virtual living room and were asked to look around and familiarize themselves with the virtual environment. Next, they were guided through a short training session in which the pursuit tracking task was introduced. Participants were encouraged to practice the task in each of the conditions, in order to get acquainted with the system and the different types of VHS shift. Once they felt confident that they could perform the task successfully in each of the conditions, the VR equipment was removed and they completed a form with demographic information.

The participants were then again equipped with the VR gear, so as to begin the trials, each of which consisted of 2 phases: In Phase 1, the illusion of ownership was elicited. The virtual and real hand were always collocated in this phase, i.e., the VHS was not shifted. Participants were asked to lift their right hand and follow a set of instructions, e.g., to look carefully at the virtual hand, wriggle their fingers and observe the virtual hand responded to their movements. This provided synchronous visuo-motor stimuli. They were then asked to extend their index finger. Then, pointing upward, a virtual ball appeared to be bouncing up and down on the tip of their index finger. Haptic feedback through the vibration motor enabled them to feel the impacts of this ball. This congruent visuo-tactile stimulation continued for about 20 seconds. Phase 2 consisted of the pursuit tracking task. The experimenter directed the participants’ attention to the virtual panel, which they would interact with during the trial. Participants were asked to reach out to the panel with the virtual hand. In the Instant shift condition, this meant that participants needed to first locate the virtual hand, since the VHS space was shifted from the start and the virtual hand was not collocated with their physical hand. In contrast, all other conditions allowed them to rely on proprioception to plan the initial reaching movement. Participants were allowed to step closer to or
further away from the panel, to accommodate variations in arm length. When
the participants indicated that they were ready to begin tracking the target
on the panel, the experimenter activated a 2-minute timer, during which their
task performance was recorded. In the *Ownershift* condition application of the
gradual shift was started at this point in time and the maximum offset was
reached after 45 seconds. In each trial, the position of the panel and the shift
of the VHS varied, depending on the condition (as shown in Table 9.1). At
full VHS shift, the vertical position offset between the real and virtual hand
averaged 65 cm, which corresponds to a rotation of 60° around the shoulder.

After each trial, the HMD was again removed and participants answered the
questionnaire (see Table 9.3). This also gave them a short break from VR and
permitted any lasting effects to wear off before the following trial. There were
5 trials, one for each condition. The order of conditions was counterbalanced
with a Latin square. Each trial (including the questionnaire) took about 5
minutes to complete, 2 minutes of which were spent on the tracking task.
This duration was based on earlier studies, which indicate that it may take
some time for the ownership illusion to occur and also to wear off. Pilot studies
were conducted to ensure that the task was not so long as to be painful.
At the end of the experiment, the participants were asked to rate their own
performance in the pursuit tracking task and were invited to talk openly about
the different conditions. The experiment took approximately 40-60 minutes.

9.6 Results

In summary, we found that the VHS shift reduced strain successfully during in-
teraction with an overhead target (H1). Furthermore, questionnaire responses
show that the illusion of body ownership persisted, despite the large offset be-
tween the real and virtual hand (H2). The gradual offset in the *Ownershift*
technique was described as more comfortable and less disruptive than the in-
stant offset in the *Instant shift* condition. Finally, based on agency ratings,
we found that participants consistently felt in control, despite a slight decrease
in task performance (<4% decrease) in the shifted conditions (*Ownershift* and
*Instant shift*) compared to the collocated conditions (*Top*, *Bottom*) (H3).

Physical strain

The amount of physical strain was evaluated from responses to questions Q5 to
Q9 of the trial questionnaire (see Table 9.3). Responses indicate that the VHS
shift successfully reduced physical strain during interactions with an overhead
target.

Significant effects of condition on strain were analyzed with one-way repeated-
measure ANOVA on questionnaire items, which were rated on the Borg-CR10
9.6. RESULTS

Figure 9.5: This plot shows the total strain per condition as a sum of mean strain ratings (on Borg-CR10 scale) for each body part. It can clearly be seen that the highest degree of strain was experienced in the Top condition, where the shoulder and upper arm were most strongly affected.

scale. Mauchly’s test was used to detect any violation of sphericity, and corrections where made when needed. Post-hoc evaluation was performed through a pairwise t-test with Bonferroni correction.

Fig 9.5 shows averages over all strain items (Q5-Q9), and reveals that the Top condition was more strenuous than all other conditions. Analyzing the questions individually, we find significant effects for all body parts (Q5-shoulder: $F(2.57, 38.50) = 22.83, p < 0.01$, partial $\eta^2 = 0.60$; Q6-upper arm: $F(4, 60) = 18.83, p < 0.01$, partial $\eta^2 = 0.56$; Q7-forearm: $F(1.90, 28.57) = 6.20, p < 0.01$, partial $\eta^2 = 0.29$; Q8-hand: $F(1.60, 24.02) = 4.35, p < 0.05$, partial $\eta^2 = 0.22$; Q9-neck: $F(4, 60) = 1.54, p < 0.05$, partial $\eta^2 = 0.09$).

Pairwise t-tests show significantly higher shoulder strain in the Top condition than in all other conditions (T vs. all others: $p < 0.01$). Reaching up to the target in the Top condition resulted in an average rating of 4.1 for the shoulder, which, according to the Borg-CR10 scale, corresponds to a perception of strain somewhere between “moderate” and “strong”. In all other conditions, average ratings for shoulder strain remained below “weak” (avg. strain $\leq 1.6$). A similar effect was evident for upper arm strain (T vs. all others: $p < 0.01$), which again averaged between “moderate” and “strong” in the Top condition ($mean = 4.3$), but remained between “very weak” and “weak” (avg. strain $\leq 1.4$) in all others. The forearm, hand, and neck were less strongly affected by the overhead tracking task with mean ratings never exceeding 2.6 (below “moderate”). We therefore omit those results from the paper.
CHAPTER 9. OWNERSHIFT PUBLICATION

Body ownership illusion

Evaluating responses to questionnaire items Q1, Q2, and Q4 (see Table 9.3), we found evidence of the body ownership illusion, despite the VHS shift in the Ownershift condition. While the findings do show that the body ownership illusion was weakened by the offset between the real and virtual hand, a comparison with the Control condition indicates that participants still perceived a significant connection to the virtual hand in the Ownershift condition.

We performed Friedman rank sum tests on the Likert scale ratings to check for significant effects, followed by post-hoc pairwise Wilcoxon signed-rank tests with continuity correction. For the main ownership question Q1 (“There were times when I felt that the virtual hand was part of my body”), we found a significant effect of condition ($\chi^2(4)=30.91, p < 0.01$). The ratings for Q1 per condition can be seen in Fig 9.6.

Post-hoc analysis revealed that ownership was significantly lower in the conditions with VHS shift (Ownershift and Instant shift condition) than in both collocated conditions, Bottom and Top. (O vs. B: $W=63, Z=2.78$, $p < 0.01, r=0.69$; O vs. T: $W=50.5, Z=2.54$, $p < 0.05, r=0.64$; I vs. B: $W=73, Z=2.63$, $p < 0.01, r=0.66$; I vs. T: $W=42, Z=2.40$, $p < 0.05, r=0.60$). There was also a significant effect between the Bottom and the Top conditions (B vs. T: $W=32.5, Z=2.16$, $p < 0.05, r=0.54$), indicating that interaction with the panel at top position reduced the body ownership illusion. However, all conditions received predominantly positive ownership ratings apart from the Control condition, for which ownership was significantly lower (C vs. O: $W=76, Z=2.98$, $p < 0.01, r=0.75$; C vs. I: $W=89, Z=3.13$, $p < 0.01, r=0.78$; C vs. T: $W=103.5, Z=3.29$, $p < 0.01, r=0.82$; C vs. B: $W=104, Z=3.30$, $p < 0.01, r=0.83$).

Responses to Q2 (“There were times when I felt like I had more than one right hand”) revealed a significant effect of condition on the illusion of owning multiple right hands ($\chi^2(4)=19.63, p < 0.01$). Pairwise comparisons show that the feeling of owning multiple hands significantly increased when there was an offset between the real and virtual hand, as previously found by Nierula et al. [92]. However, responses were predominantly negative.

Agency ratings

The feeling of agency over the virtual hand was evaluated from item Q3 (“There were times when I felt I could control the virtual hand as if it were my own.”). As can be seen from Fig 9.7, participants gave positive ratings in all conditions, apart from the Control condition, indicating that they felt in control, despite the VHS shift.

The results of a Friedman rank sum test show a significant effect of condition on agency ($\chi^2(4)=34.68, p < 0.01$). Post-hoc pairwise Wilcoxon signed-
9.6. RESULTS

Figure 9.6: Responses to the primary ownership item (Q1) were highest in the Bottom condition and significantly lower in conditions with VHS shift (Ownershift and Instant shift condition). However, all synchronous conditions yield evidence of body ownership, with only the Control condition prompting mostly disagreement. Significant effects are indicated by * (p < 0.05) and ** (p < 0.01).

Task performance

Analyzing participant performance during the tracking task, we observed a decrease in task performance in both shift conditions (Ownershift, Instant shift), compared to the Top and Bottom conditions (see Fig 9.8). Generally, we found that errors were slightly larger when interacting with the panel at the top, even without a shift (Top condition). As expected, the largest errors were recorded in the Control condition.

Mauchly’s test showed a violation of sphericity against Condition (W(4) = 0.03, p < 0.05). With a corrected one-way repeated-measure ANOVA, we found a significant effect of condition on the error ratio (F(1.61,24.1) = 156.86, p < 0.01, partial $\eta^2 = 0.91$).

Post-hoc analysis using a pairwise t-test with Bonferroni correction reveals the following effects: the VHS shift in Ownershift and Instant shift condition led to significantly higher errors than both collocated conditions, Bottom and
Figure 9.7: Participants perceived a high degree of control over the collocated virtual hand \((\text{Bottom, Top})\). Agency remained similarly strong, despite a shift of the VHS, as long as the hand motion remained synchronous \((\text{Ownershift, Instant shift})\). As expected, the asynchronous \textit{Control} disrupted the feeling of agency. Significant effects are indicated by * \((p < 0.05)\) and ** \((p < 0.01)\).

\textit{Top} \((\text{O,I vs. B,T: } p < 0.01)\). However, in interaction with the panel at the top, average errors were similar \((\text{O: } \text{avg.error} = 0.199, \text{ T: } \text{avg.error} = 0.16)\) and the increase in error ratio between the \textit{Ownershift} condition and \textit{Top} condition amounted to less than 4\%. There was no significant difference between \textit{Ownershift} and \textit{Instant shift} \((\text{I vs. O: } p = 0.75; \text{ O: } \text{avg.error} = 0.199, \text{ I: } \text{avg.error} = 0.197)\). When hands were collocated, interaction with the panel at the top also led to a larger task error than with the panel at the bottom \((\text{B vs. T: } p < 0.01; \text{ B: } \text{avg.error} = 0.14, \text{ T: } \text{avg.error} = 0.16)\). Finally, as expected, participants displayed the lowest overall task performance in the \textit{Control} condition \((\text{C vs. all others: } p < 0.01)\) with 43\% error \((\text{avg.error} = 0.43)\).

In the \textit{Ownershift}, the VHS shift was increased linearly over 45 seconds during the first minute of interaction. Earlier research shows evidence that a drifting hand position may lead to reduced performance in pointing tasks [19]. However, we found no significant differences when comparing the error ratios of the first and second minutes of the task, in order to explore whether task performance was diminished when participants needed to compensate for the changing offset.

\textbf{Post-trial questionnaire and qualitative findings}

At the end of the experiment, participants rated their own task performance on a 7-point Likert scale (from -3 ‘strongly disagree’ to 3 ‘strongly agree’). They were then asked to comment on each of the conditions in a semi-structured interview. Participants were encouraged to say the first thing that came to
9.6. RESULTS

Figure 9.8: Task performance, measured as the ratio of the root-mean-square error to the target movement, was best in the Bottom condition. Interaction with the top targets led to a higher error rate, in particular when a VHS shift was present (Ownershift, Instant shift). Errors in the Control condition, however, were nearly twice as large and performance was thereby significantly worse than in all other conditions. Significant effects are indicated by * ($p < 0.05$) and ** ($p < 0.01$).

Our findings show that overall, participants felt good about their task performance. They felt they could adjust quickly to the VHS shift in the Ownershift and the Instant shift conditions, and that interaction with the panel worked well. While the Instant shift required a short “calibration phase”, most participants were not very much aware of the shift in the Ownershift condition. Both shift conditions were perceived as much more comfortable than the Top condition. As expected, the Bottom condition felt most natural, while the Control condition was perceived as annoying and disruptive.

Self-assessment of task performance

Participants felt that they had performed fairly well in the pursuit task. The statement “I did well in the pursuit tracking task (i.e., following the dot on the panel)” received a median rating of 1.5 (on a scale from -3 to 3).

Qualitative findings for interaction conditions

Participants were asked to comment on each of the conditions in the order in which they had experienced them. Note that participants were not aware that the Ownershift condition was the technique we wished to evaluate.
Our interaction technique, Ownershift, was described as “natural” (p5, p7), and participants indicated that hand movements were predictable (p6, p12) and that interaction became more comfortable as the shift occurred (p4, p13). Two participants described the interaction as a little “strange” (p10, p15). The majority of participants (12/16) indicated that they were only slightly or not at all aware of the shift taking place. The adjustment to the offset happened “automatically” (p7, p9) and some recounted that they suddenly became aware of their real hand being much lower than the virtual hand (p3, p10). Only two participants (p2, p13) noticed the shift strongly (p13 - “especially in big circular motions”). The gradual offset in the Ownershift condition was often described as more comfortable and easier to adapt to than the instant offset. The gradual shift was described as requiring less cognitive load and being less disruptive to their connection with the virtual hand (p1, p3, p5, p7, p16). Participants said “it felt more like my own hand than with the instant offset, because I didn’t notice the difference as much”(p1), and “[the Ownershift condition] feels a bit more natural than instant. […] there’s a more natural connection.” (p7). Another participant added that Ownershift “wasn’t as jarring as when my arm was initially up there [referring to the Instant shift condition]” (p16). However, one participant explicitly indicated that she preferred the Instant shift condition, because there “you had a clear idea of what your movement should be […] when it was gradually offset, you had to adjust all the time and that was a bit difficult” (p2).

The Instant shift condition was described as “nice” and “better”, “more convenient”, “more comfortable” and “less tiring” than Top (p2, p7, p9, p10, p12, p13, p16). Participants indicated that interaction worked well, did not require much concentration, and that recalibration was automatic (p1, p3, p7, p10, p11, p13, p14, p15, p16). Interestingly, many explicitly described a first calibration phase and said that once they had made the mapping in their brain, it was similar to interacting with collocated hands (p3, p8, p12, p13, p14, p15): “there was a calibration phase but when I got it, it was fine” (p12); “when you learn it, you learn it” (p3). Others commented that the virtual hand felt more disconnected and less like their own hand (p4, p6) in the Instant shift condition. They described it as “remote controlling” something (p7); “it was (as if) connected with a string or […] balanced the hand on top of my hand” (p4); “it was like projecting your hand up there […] like a laser pointer” (p5).

Unsurprisingly, the Top condition, with collocated hands and interaction overhead, was described by all participants as “tiring”, “uncomfortable”, “painful” or “annoying”. For instance p11 said “I couldn’t last more than 2 minutes”. One participant (p3) said that he felt like time went much faster in the trials with VHS shift, and that the Top condition was the longest trial. But some
participants also described it as “good” (p10, p11) in terms of control, “real” (p5) and similar to Bottom (p13, p14, p15).

9.7 Discussion

From our findings, we conclude that a vertical shift (65 cm) of the VHS can effectively reduce fatigue during overhead interaction (H1), while maintaining the illusion of owning the virtual hand (H2). Performance is only slightly affected by this shift (less than 4% decrease) and the user feels in control of the virtual hand (H3). The decrease in task performance could be addressed by increasing the size of the moving target; with a target of 3 cm radius, participants would have stayed within the target’s bounds 90% of the time (in the Ownershift condition, 90% of measured errors were smaller than 0.03). Additionally, it stands to reason that training could improve task performance [45]. As interpreted by Kohli et al. [68], in a warped virtual space, users are less precise at touching targets, but precise enough. However, it remains to be explored how a gradual shift might affect more delicate manipulations (e.g., drawing a straight horizontal line).

We find that participants performed equally well with the gradual shift in the Ownershift technique, as when interacting with an instantly shifted virtual hand, which has earlier been shown to be more effective than retargeting through interpolation [45]. The qualitative results further suggest that a gradual shift is preferable, since it was described as much less noticeable and less disruptive. Furthermore, it did not require a mental recalibration phase, since the initial 1:1 mapping in Ownershift allows initial ballistic movements toward the target (RQ1). This is in line with earlier work [45], which found that the instant shift was disorienting, when it meant that participants first had to look around to locate the hand, before being able to start interacting. Importantly, we found no evidence that task performance was diminished, while the VHS was gradually being shifted (RQ2).

We can confirm that a position mismatch between the real and the surrogate hand (i.e., rubber hand or virtual hand) leads to lower ratings for the body ownership illusion, as has previously been found by several authors [60, 75, 92]. However, our findings also indicate that under certain circumstances (i.e., realistic hand representation with a combination of visuo-motor and visuo-tactile stimulation), the VIH is much more robust towards such position offsets than was previously believed. It is conceivable that the illusion of owning a virtual hand could occur with even more extreme shifts. However, this remains to be explored. Surprisingly, we observed lower ownership in interaction with the panel at the top, even when hands were collocated (Top condition), compared to interaction with the panel at the bottom (Bottom condition). Presently, we can only surmise about the reasons for this. For instance, it may simply be due to less accurate hand tracking in that condition, since the LeapMotion is
not optimized for interaction with elevated hands; or it could even be a side-effect of the discomfort of interacting overhead. More research is required to verify these assumptions.

In future, to make the shift even less noticeable, we recommend applying the gradual shift in proportion to the user’s own hand motion. The user’s movement could thus completely mask the displacement of the virtual hand, as has been explored in redirected walking [101] and haptic retargeting [8]. Furthermore, it remains for future work to explore the maximum unnoticeable VHS shift speed, as well as to devise better approaches for resetting the shift and dynamic switching between targets.

Apart from the aim of reducing fatigue, a technique based on Ownershift could also pursue support people with reduced mobility. By adjusting the shifts to their specific limitations, it could allow them to access the full reachable space. Furthermore, the principle could be inverted to intentionally increase the difficulty of reaching tasks for rehabilitation purposes.

The pursuit tracking task [98] proved well suited for the evaluation of continuous interaction with a target in VR. It kept participants engaged and required smooth hand movements at moderate speeds, which resembles prolonged direct manipulation of a virtual object. This was a common task for evaluating user interfaces in the 70’s, but has largely been replaced by Fitts’ law tasks [79] in contemporary HCI research. While tracking tasks do not allow measuring the throughput of a user interface (in bit/s), they evaluate how well users can control a dynamic system. Therefore, we believe that for interaction without a natural delimiter, such as free-hand interaction in immersive VR, the pursuit tracking task provides valuable insights. This is especially true if we wish to evaluate dynamics control, such as the ability to interact with physical simulations, as opposed to mere button clicking. Through this argument, we hope to inspire more researchers in HCI to use tracking tasks for evaluating user interfaces, in addition to Fitts’ law tasks.

9.8 Conclusion

This paper presents Ownershift, an interaction technique that reduces strain during the manipulation of overhead targets, while preserving the illusion of owning the virtual hand. This is achieved by first supporting fast reaching for different targets with a 1:1 mapping, and then gradually shifting the virtual hand space (VHS) upwards, towards an overhead virtual target for prolonged interaction. This gently directs the user to move his hand into a more comfortable position at waist-level. While this leads to a radical offset between the real and the virtual hand, we found that users adjust to this easily, and that the impact on their task performance is limited (less than 4% increase in error). Our user study also provides evidence that a gradual application
9.9 APPENDIX

of VHS shift is preferable to an instant shift, since it allows initial ballistic reaching towards a target, and the offset is easier to adapt to and causes less ‘disconnection’ from the virtual hand. Furthermore, in contrast to earlier work, we found that the body ownership illusion can be maintained, despite the large vertical position mismatch (avg. 65 cm) between the real and the virtual hand.

9.9 Appendix

Design of the pursuit tracking task

In the user study, task performance was evaluated with a pursuit tracking task [98] in which the participant was instructed to follow a moving target with his finger as closely as possible. The quasi-random movement of the target results from the sum of 4 sinusoids for both the $x$ and $y$ directions. Table 9.1 contains the angular frequency and phase shift for each sinusoid.

<table>
<thead>
<tr>
<th>axis</th>
<th>sin1</th>
<th>sin2</th>
<th>sin3</th>
<th>sin4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>$2.07t$</td>
<td>$1.65t$</td>
<td>$0.23t$</td>
<td>$0.37t$</td>
</tr>
<tr>
<td>$y$</td>
<td>$2.07t + 0.5\pi$</td>
<td>$0.96t + 0.5\pi$</td>
<td>$1.98t$</td>
<td>$1.08t$</td>
</tr>
</tbody>
</table>

Table 9.1: Angular frequency and phase shift of the quasi-random target motion in the pursuit task. The amplitude was 1.2 and $t$ denotes the time in seconds.

Quasi-random VHS shift

The purpose of the Control condition was to break the body ownership illusion, reduce the feeling of agency and provide a benchmark for unacceptable task performance. Therefore, a quasi-random VHS shift was applied, which resulted in a constantly changing offset between the virtual hand and real hand positions, consequently rendering the hand uncontrollable for the user. The quasi-random shift was calculated as a sum of sinusoids, as presented in Table 9.2.

<table>
<thead>
<tr>
<th>axis</th>
<th>sin1</th>
<th>sin2</th>
<th>sin3</th>
<th>sin4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>$2.51t$</td>
<td>$2.4t$</td>
<td>$0.28t$</td>
<td>$0.44t$</td>
</tr>
<tr>
<td>$y$</td>
<td>$2.51t$</td>
<td>$0.12t + 0.25\pi$</td>
<td>$2t$</td>
<td>$2.51t + 0.5\pi$</td>
</tr>
<tr>
<td>$z$</td>
<td>$1.31t$</td>
<td>$1.16t + 0.5\pi$</td>
<td>$2.4t$</td>
<td>$0$</td>
</tr>
</tbody>
</table>

Table 9.2: Angular frequencies and phase shifts of the quasi-random hand offsets applied in Control condition. ($t$ is time in seconds)
CHAPTER 9. OWNERSHIFT PUBLICATION

Tracking task error

During the pursuit tracking task, the task error was calculated as the root-
mean-squared error (\( \text{rmse} \)) [98]:

\[
\text{rmse} = \sqrt{\frac{\sum_{t=1}^{n} |P_F(t) - P_T(t)|^2}{n}},
\]

(9.1)

where \( P_F \) is the position of the participant’s fingertip and \( P_T \) is the target position at time \( t \).

The participant’s task performance was then computed as the ratio of task error (\( \text{rmse} \)) to a reference error (\( \text{refE} \)) that would occur if the participant’s finger had remained in the center of the panel during the entire trial:

\[
\text{taskperformance} = \frac{\text{rmse}}{\text{refE}}
\]

(9.2)

with

\[
\text{refE} = \sqrt{\frac{\sum_{t=1}^{n} |C - P_T(t)|^2}{n}},
\]

(9.3)

where \( C \) is the center of the panel.

Questionnaire

After each trial in the user study, participants were asked to respond to the questionnaire items in Table 9.3. The questionnaire was completed on paper.

Body ownership & agency

Q1 There were times when I felt that the virtual hand was part of my body.
Q2 There were times when I felt like I had more than one right hand.
Q3 There were times when I felt I could control the virtual hand as if it was my own.
Q4 I never felt that the virtual hand was part of my body.

Physical strain

Q5 How much strain did you feel in your right shoulder?
Q6 How much strain did you feel in your right upper arm?
Q7 How much strain did you feel in your right forearm?
Q8 How much strain did you feel in your right hand?
Q9 How much strain did you feel in your neck?

Table 9.3: Participants responded to this questionnaire after each trial, indicating their degree of ownership (Q1, Q2, Q4), agency (Q3) and physical strain (Q5-Q9) for each condition.
Exploring different Multisensory Stimuli and Visual Realism for the Virtual Hand Illusion

Tiare Feuchtner¹, Jörg Müller² and Mel Slater³

¹Department of Computer Science, Aarhus University, Denmark
²Department of Computer Science, University of Bayreuth, Germany
³Department of Clinical Psychology and Psychobiology, University of Barcelona, Spain

Abstract

We elicit the virtual hand illusion through different multisensory stimuli and explore how these modulate body ownership of realistic and unrealistic hands. We compare visual-motor, visual-proprioceptive and cardio-visual stimuli to elicit body ownership of three different virtual hand representations: a very realistic hand, an abstract hand model, and a very unrealistic, disconnected hand representation. For visual-motor synchrony we support full hand tracking, allowing natural movement of the virtual hand. The visual-proprioceptive and cardio-visual conditions are so-called vision-only conditions (no hand tracking). In the latter the participant’s heartbeat is visualized on the virtual hand. Our results show that realism of the virtual hand significantly affects the body ownership illusion. We found no significant differences between the vision-only and motor conditions, nor was there an influence of the heartbeat visualization. However, 50% of participants indicated body ownership even of very unrealistic hands, if they could move the hand.

CCS Concepts

• Human-centered computing → Virtual reality; Empirical studies in HCI;

Additional Key Words and Phrases

body ownership illusion; virtual hand illusion; virtual reality; visuo-motor; cardio-visual.

10.1 Introduction

This paper reports an experimental study, which aims to evaluate the impact of different multisensory stimuli and visual hand representations on the virtual hand illusion (i.e., the feeling of body ownership of virtual hand).

Most experiments that study the illusion of ownership of an artificial hand rely on synchronous visual-tactile stimulation of the real and surrogate hand. For instance, in the original rubber hand illusion experiment [15], a participant’s real hand and a rubber hand were synchronously stroked with a brush, while the real hand was hidden from view and only the rubber hand was visible.
CHAPTER 10. PHYSIOMOTOR PUBLICATION

Feeling brush strokes that coincide with those observed on a rubber hand leads to the interpretation that touch can be felt on the rubber hand and therefore it must be part of one’s body. Such ownership illusions can also involve virtual limbs, for instance when experiencing a virtual hand as the own in an immersive virtual environment [105]. It has been shown that the illusion of ownership can even be achieved without tactile feedback, e.g., through visual-proprioceptive synchrony when simply observing a virtual hand in place of the own (“vision-only”), or visual-motor synchrony where the virtual hand moves in synchrony with the physical hand. Furthermore, the work of Blanke et al. [7] suggests that feedback of body-internal processes, such as cardiovisual feedback, may have a positive impact on embodiment. In this study we explore how such different combinations of multisensory stimuli affect the illusion of owning a virtual hand.

Further, it has been found that the appearance of the artificial hand affects the prevalence of the ownership illusion. Experiments have shown that ownership is stronger for realistic hand representations than abstract ones [6, 73]. There is evidence that a limb can only be perceived as part of the body if plausible morphology and placement are given [25, 127, 128] and the artificial hand appears connected to the body [95, 123, 124]. However, it remains to be explored whether different combinations of multisensory stimuli might be successful in overcoming limitations of unrealistic hand representations. We hypothesize in particular that visual-motor synchrony could lead to ownership of disconnected hands, since disconnected elements moving together can be perceived as being connected. In other words, when perceiving the movement of a hand consisting of disconnected parts the human brain might fill in the gaps resulting in a coherent hand representation. We explore this premise by presenting participants with different hand representations with varying degree of realism and connectivity.

In a mixed 3x3 experimental design we varied three Stimulation Types between-groups (Motor Condition, Physio Condition, Passive Condition), and three Visual Appearances as within-groups factor (hand A, hand B, hand C). While Motor Condition provided visual-motor synchrony through hand motion tracking (i.e., the virtual hand mimicked the movements of the participant’s physical hand), participants could only observe an immobile virtual hand in place of their own hand in both Physio Condition and Passive Condition (visualproprioceptive synchrony). In addition, in Physio Condition participants perceived their own heartbeat visualized on the virtual hand as a subtle change of color. The virtual hand representation was either a realistic-looking human hand (A), an abstract skeleton-like hand (B), or an unrealistic hand with claws for finger tips and no visible palm or fingers (C).

Our results show that the type of multisensory stimuli has surprisingly little impact, with vision-only (Passive Condition and Physio Condition)
achieving similar ownership ratings as visual-motor synchrony (Motor Condition). Furthermore, we found no effect of visualizing internal signals on self-identification (Physio Condition) [7]: despite noticing the color change, participants did not recognize the flashing as their heartbeat and ownership ratings were similar to the other conditions. Furthermore, we did not find a correlation of interoceptive sensitivity and the occurrence of the ownership illusion [120, 129]. We did however observe that the visual representation of the virtual hand strongly impacts the ownership illusion: it is more likely to experience ownership of a realistic-looking virtual hand than an unrealistic one, which is in line with related work [6, 73]. We can also confirm the earlier observation that missing connectivity of the virtual hand leads to a decrease in ownership [95, 123, 124]. Nevertheless, since 50% of participants gave positive ownership ratings for the unrealistic hand in the Motor Condition, visual-motor synchrony may to some degree compensate incongruities like missing connectivity.

10.2 Background

There is abundant evidence showing that our internal body representation is flexible and constantly updated based on processed multisensory information. This has for instance been studied through body ownership illusions (BOIs) [11, 50, 66, 96], which involve the self-attriubtion of external objects to the own body. Such illusions can be induced by providing congruent multisensory information [20, 80], for instance through synchronous visual-tactile stimulation of an artificial body and its real counterpart. Notably, these illusions may concern single artificial limbs [15, 43, 64, 115], or even of a whole artificial body [81, 96, 117].

The probably most widely known BOI is the Rubber Hand Illusion (RHI), discovered by Botvinick and Cohen [15]. In the classical RHI paradigm a rubber hand is placed on a table next to the participant’s hidden real hand and both hands are simultaneously stroked with brushes. Since the tactile stimuli the participant feels coincide with the brush strokes seen on the rubber hand, this leads to the illusion that the rubber hand is in fact part of the participant’s body. Interestingly the illusion does not appear to depend on tactile stimulus. It can also be elicited through visual-proprioceptive or visual-motor correlations [26, 32, 59, 105, 127, 132] when experiencing an artificial body (part) from first person perspective in place of the own body (part).

The artificial body can either be a mannequin, or a virtual avatar seen in virtual reality (VR). Immersive VR provides a very high degree of control over the participant’s environment and body representation. For instance, we can drastically transform a user’s virtual arm, changing it in size [34, 64], connectivity [95, 123], or appearance [6, 73]. Hence it has become a popular tool for exploring the ownership illusion, where the feeling of owning a virtual
CHAPTER 10. PHYSIOMOTOR PUBLICATION

hand has been dubbed the virtual arm illusion [115], or virtual hand illusion (VHI) [105].

Variations of RHI and VHI experiments were conducted to explore differences in multisensory stimuli, hand position and orientation [25, 95], as well as identity and appearance of the artificial limb [6, 73, 95, 123, 127, 128]. The results reveal some factors to be critical for the ownership illusion, such as matching stimuli, as well as congruent orientation, identity, and connectivity of the surrogate body part. It is in particular the aspect of appearance and connectivity of an artificial hand that we aim to explore in this paper.

Ownership illusions have mainly been studied based on exteroceptive signals. However, it has been found that our body representation stems from the integration of both interoceptive and exteroceptive signals [7]. Physiological signals, like the sound of breathing or our heartbeat, is something that we associate strongly with our own body [130]. Groenegress et al. [42] explored the metaphorical visualization of physiological measurements, such as heartbeat and breathing, and it was found that synchronous cardio-visual signals can increase the illusion of ownership of a virtual body [7] or body part [120]. Further research has shown that our degree of interoceptive sensitivity is a predictor for the malleability of our internal body representation: people with lower interoceptive awareness tend to experience stronger ownership illusions [129]. Conversely it has also been found that body ownership illusions can modulate interoceptive sensitivity [37]. For people with low interoceptive sensitivity, experiencing a body ownership illusion can lead to an increased awareness of body-internal processes.

Our work aims at exploring how different multisensory stimuli (i.e., visual-proprrioceptive, visual-motor and cardio-visual stimuli) affect the VHI for unrealistic hand representations. Prior research has focused on each of these aspects separately, but not in combination. It seems plausible that some combinations of stimuli fare better at compensating for visual inconsistencies between the real and virtual hand, and it is for the purpose of testing this assumption that we conducted the experiment described in this paper.

10.3 Experimental design

To explore the impact of different multisensory stimuli on visual incongruities between the real hand and virtual hand, we conducted an experiment with two independent factors: Stimulation Type and Visual Appearance.

The following three Stimulation Types (conditions) stimulation types were varied between groups (rows of Fig 10.1).

1. **Motor Condition:** the virtual hand moved in synchrony with the movements of the participant’s real hand and fingers, providing *visual-motor*
10.3. EXPERIMENTAL DESIGN

Figure 10.1: **Combinations of hand representations and multisensory stimuli.** This matrix shows the independent variables of our study: Visual Appearance of the virtual hand (columns) and Stimulation Type (rows). Each participant experienced one Stimulation Type (between groups). From top to bottom: in **Motor Condition** the virtual hand follows the participant’s movement, (showing a participant making a “thumbs-up” gesture); in **Physio Condition** participants could see the unmoving virtual hand pulse in synchrony with their heartbeat (visualized by the hand turning red and then fading to skin color with each beat); in **Passive Condition** participants looked at an unmoving virtual hand in similar pose and location to their own hand without any additional feedback. Furthermore, each participant experienced all levels of Visual Appearance (within groups), one after the other. From left to right: most realistic hand with detailed mesh (A); abstract hand with skeleton-like fingers (B); least realistic hand consisting only of fingertips and the stump of the arm (C).

2. **Passive Condition:** the virtual hand was seen resting on a desk in similar position and posture as the participant’s real hand (*visual-proprioceptive synchrony*).

3. **Physio Condition:** this condition was the same as the Passive Condition, but in addition the virtual hand provided physiological feedback by subtly pulsing (changing color) in synchrony with the participant’s heartbeat (*visual-proprioceptive and cardio-visual synchrony*).
CHAPTER 10. PHYSIOMOTOR PUBLICATION

In the latter two conditions participants were asked to keep their right hand still and hand movement tracking was deactivated.

The types of Visual Appearance (hand representations), which were varied within groups, were as follows (columns of Fig 10.1).

1. **Hand A**: the virtual hand had a detailed and realistic visual representation, including creases and fingernails.

2. **Hand B**: the hand model was simplified, blocky, and the fingers remained separated all the way to the wrist, as in a skeleton (i.e., there was no connected palm).

3. **Hand C**: the palm and base of the fingers were invisible, so there was no apparent connection between the claw-like fingertips and the wrist.

We had a mixed 3x3 experimental design, with Stimulation Type as between-groups factor, and the Visual Appearance as within-groups factor: the participants were divided into three groups, so that each participant experienced one type of multisensory stimulation and all virtual hand representations. The different virtual hands were presented in three consecutive trials, either in increasing or decreasing order of realism (i.e., t1:A - t2:B - t3:C or t1:C - t2:B - t3:A). After each trial the participants responded to a series of questions indicating the degree of ownership they felt of the virtual hand, as is explained in more detail in the following section. To be able to rate the hands comparatively, participants were shown each of the virtual hand representations before the trials.

Our first hypothesis was that ownership would diminish along with the hand’s realism. Thus, we expected a high degree of ownership for the most realistic hand representation (A) and ownership to be low or non-existent for the least realistic virtual hand (C) (H1). However, we hypothesized that visual-motor synchrony (Motor Condition) could elicit such a strong illusion of ownership that incongruities in the visual appearance of the virtual hand (i.e., unrealistic appearance, missing palm, no connectivity of fingers to arm) would be tolerated (H2). Furthermore, we predicted that visualizing physiological signals on the virtual arm, such as a subtle color change in the rhythm of the heartbeat (Physio Condition), could increase the illusion of ownership in absence of visual-motor synchrony (H3).
10.4 Materials and methods

Apparatus

The experiment was implemented in the Unity3D\(^1\) game engine. A head mounted display (HTC Vive\(^2\)) allowed the participant to become immersed in the virtual environment. A LeapMotion\(^3\) sensor was attached to the front of the head mounted display to track the participant’s hand and finger movements, as is shown in Fig 10.2. The participants interacted with the system through a PowerMate\(^4\) controller which they controlled with their left hand. For instance, they could rotate the button to indicate the virtual hand position and select questionnaire responses, and the button could be pressed down to confirm their input and continue with the study.

The ECG signal was sampled at 256Hz using a g.USBamp bio-signal amplifier by g.tec\(^5\) and pre-processed with Simulink\(^6\). The filtered data was then recorded and streamed to Unity to give cardio-visual feedback in the Physio Condition, as explained below.

Implementation of the heartbeat visualization.

For the visualization of the heartbeat we implemented a subtle change of color of the virtual hand, based on implementations reported in earlier work [7, 120]. Each heartbeat results in a visible spike in the ECG signal (QRS complex) [42] and each such spike was detected using a simple threshold. This threshold was calibrated for each participant to guarantee reliable heartbeat detection.

---

\(^1\)https://unity3d.com
\(^2\)https://www.vive.com
\(^3\)http://leapmotion.com/
\(^4\)https://griffintechnology.com/us/powermate
\(^5\)http://gtec.at
\(^6\)https://se.mathworks.com/products/simulink.html
CHAPTER 10. PHYSIOMOTOR PUBLICATION

Figure 10.3: Heartbeat visualization. In the Physio Condition participants could see the virtual hand changing color in accordance with their own heartbeat. Using a threshold (green line) we detected each spike in the raw physio data at which we set the hand’s color to red. It then slowly faded back to its original color, before the next heartbeat.

At each detected spike the color of the virtual hand was immediately changed to dark red and then slowly faded back to the normal skin tone before the next heartbeat (see Fig 10.3). This resulted in a pulsing effect, as if the blood ejected by the heart was briefly coloring the hand red.

Measurements
The body ownership illusion was evaluated mainly based on an adaptation of the RHI questionnaire, as explained below. Additionally, we measured proprioceptive drift to corroborate our subjective data. Furthermore, interoceptive awareness was evaluated based on a heartbeat counting task (mental tracking method) and a body awareness questionnaire described in the following sections.

Body ownership questionnaire
To evaluate the perceived level of body ownership we adapted questionnaires used in earlier work [15, 59, 64, 115] resulting in items Q1 to Q4 in Table 10.1. Participants rated these items interactively after each hand trial and in randomized order (an example is shown in Fig 10.4). The main ownership question, Q1, refers to the participant perceiving the virtual hand as part of the own body. Q2 is a negated variant of Q1, thus ratings are expected to be low when they are high for Q1 and vice versa. Q3 is considered a control question, since the VHI or RHI usually leads to a “replacement” of the real hand.
10.4. MATERIALS AND METHODS

Figure 10.4: Interactive Questionnaire. After each hand trial participants were asked to respond to a series of statements. Ratings were given by rotating a button to position the blue marker, and confirming with a button press.

by the artificial hand, instead of feeling like owning an additional right hand. The fourth item, Q4, measures the feeling of agency over the virtual hand. This question was phrased in a way that it could apply to participants in the vision-only conditions (Physio Condition and Passive Condition), who were asked to only imagine how the virtual hand would respond if they moved their hand, as well as participants in Motor Condition who were instructed to move their hands and could observe the virtual hand following their movements.

Proprioceptive drift in a position estimation task

Under normal conditions humans can give a fairly good estimate of their hand position, even without looking at it. However, when under the influence of the RHI, the estimated hand position will drift towards the position of the rubber hand [15].

To be able to measure such an effect we refrained from completely collocating the real and virtual hands. Instead, we positioned the virtual hand next to the participant’s real hand with an offset of 10 cm towards their mid-line. Hence, we expect to see a proprioceptive drift to the left, if the VHI occurs. To record the participants’ perceived hand position we hid the virtual hand and showed ruler markings on the virtual tabletop (see Fig 10.5). We then asked the participants to imagine a straight line along their index finger and to indicate the position where this line would cross the ruler. The indication was made by moving a blue marker to the respective position. This task was repeated twice, once with the blue marker appearing at the left end of the ruler and once at the right end. The order of starting positions was counterbalanced.

To prevent participants from choosing remembered positions for the marker, the desk was kept featureless and the markings on the ruler were not numbered. Participants were reminded to keep their right hand motionless during
CHAPTER 10. PHYSIOMOTOR PUBLICATION

Hand trial questions
Q1 I felt as if the virtual hand might be part of my body.
Q2 I felt like the virtual hand was NOT part of my body.
Q3 I felt as if I had more than one right hand.
Q4 I felt like I might be able to control the virtual hand as if it was part of my own body.

Post-trial questions
Q5 There were times when I noticed that my virtual hand was changing color/flashling.
Q6 It seemed as if the flashing of the virtual hand was my heartbeat.
Q7 It seemed as if I was feeling my heartbeat in the virtual hand.
Q8 It was easy for me to count my heartbeats.
Q9 I did well in the heartbeat counting task.
Q10 Which hand do you prefer? Why?

Table 10.1: Adapted ownership questionnaire. This questionnaire was used to evaluate ownership, agency and awareness of visualized physiological feedback. All items apart from Q10 were to be rated on a 7-point Likert scale. Q1 to Q4 were rated interactively in VR after each hand trial.

Figure 10.5: Measurement of proprioceptive drift. Participants were asked to indicate the position of their real hand by moving a marker along a ruler that was visualized on the top of the table. Proprioceptive drift is measured as the error in the position estimate.

the task.

Interoceptive sensitivity
To establish a continuous measure of interoceptive sensitivity (IS), we implemented a mental tracking method [106, 129]. At the beginning of the exper-
10.4. MATERIALS AND METHODS

Participants were asked to concentrate on their own body and count their heartbeats as best as they could. After a practice trial there were three counting trials of varying duration (20 seconds, 30 seconds and 55 seconds), which were presented in fully counterbalanced order. Each counting trial began with the word ‘go’ and ended with the instruction ‘stop’. The participants were then asked to indicate their heartbeat-count by turning the knob on the controller and confirm their input with a button press. During this task participants were not allowed to take their pulse, received no indication as to the duration of each trial and were not informed about their performance. The aim of these tasks was to get the participants to focus on their own physiological signals and to evaluate their IS based on their awareness of their own heartbeat.

The IS for each participant was calculated as the ratio of counted to recorded heartbeats [129]:

\[
IS = \frac{1}{3} \sum (1 - \frac{|\text{recorded} - \text{counted}|}{\text{recorded}})
\]

This results in a number between 0 and 1, where a higher number signifies less error and therefore, higher sensitivity.

After the experiment, participants responded to questions 8 and 9 (Table 10.1), which served to evaluate the participants’ own assessment of their interoceptive awareness.

**Body awareness questionnaire**

As second indicator of their awareness of bodily processes, participants were asked to complete the body awareness questionnaire (BAQ) [110]. The 18 items on the BAQ were rated on a 7-point Likert scale (1: “not at all true of me”, 7: “very true of me”). These ratings were summed and averaged to calculate a body awareness index per participant (between 0 and 1), where a higher value corresponds to higher body awareness.

**Effectiveness of cardio-visual feedback**

To assure that the cardio-visual feedback was effective, participants responded to four additional questions at the end of the experiment (Q5-Q9, Table 10.1). These questions were loosely based on the questionnaire used by Aspell et al. [7]. The first two items (Q5 to Q6) served to evaluate whether participants were consciously aware of the subtle color change on the virtual hand (only present in the *Physio Condition*), and whether they recognized their own heartbeat in it. Q7 queried whether the color change made participants feel their own heartbeat in the virtual hand. The order in which these questions appeared was randomized.
CHAPTER 10. PHYSIOMOTOR PUBLICATION

Figure 10.6: Hand representations of varying realism. These hand representations were subjectively rated by realism in an online survey (Experiment 1). The degree of realism of the hand representations declines from left to right. Three hand representations, outlined in yellow (Visual Appearance types A, B and C), were chosen by general agreement to ensure that they were clearly distinct from each other.

Preference of Visual Appearance

At the end of the study participants were asked to indicate their preference for the virtual hand representations in three pairwise comparisons (A-B, A-C, B-C). Item Q10 in Table 10.1 was repeated three times, showing pairs of virtual hand representations that were to be compared. The order in which these pairs were presented was randomized.

User study

Survey on hand representations

To explore the effect of Visual Appearance on the VHI, it was important to find three hand representations that were clearly perceived as different in respect to their degree of realism. For this purpose, we conducted an online survey in which participants were asked to rate five different hand representations (Fig 10.6), by choosing the hand which they perceived as more realistic in 10 pairwise comparisons (every hand representation was compared to each of the 4 other hands). An example of a pairwise comparison presented by the survey can be seen in Fig 10.7). The order in which the hand pairs were presented was randomized.

The survey was distributed by email to researchers at three labs (the eventLab in Barcelona, as well as the UBI and CMA groups at Aarhus University), all of which work in the fields of psychology and human computer interaction. The
10.4. MATERIALS AND METHODS

Figure 10.7: **Pairwise comparison of hand representations.** Example taken from the online survey, in which participants were asked to indicate the hand they perceived as more realistic.

survey was made available for the duration of 7 days during which responses were recorded from 28 participants (11 female), aged 22 to 57 (mean = 33.54, sd = 8.33).

From the survey we obtained the following scores per hand type: hand301 - 111 votes, hand302 - 85 votes, hand303 - 47, hand304 - 35, hand305 - 2. For clarification, if all participants had consistently rated hand301 as more realistic in comparison to hand302, hand303, hand304 and hand305, it would have collected a total of 112 votes (28 participants * 4 comparisons = 112). Thus, with 111 votes hand301 was clearly perceived as most realistic. hand302 ranks second, hand303 and hand304 were somewhat closer together and hand305 was very clearly perceived as the least realistic hand representation.

We then selected three hands with distinct levels of realism by choosing the most realistic, least realistic, and a hand that was perceived to be in the middle between those two. These hand representations, marked in Fig 10.6 with yellow outlines, included a detailed human hand (Hand A), an abstract skeleton-like hand (Hand B), and a hand without visible palm or fingers (Hand C).

**Participants**

We invited 66 naive participants to take part in the study. They were recruited through emails and fliers that were distributed throughout the buildings of the Department of Psychology at the Universitat de Barcelona. Most participants were therefore students or professionals in the field of psychology (80%).

Participants were not admitted to the experiment if they indicated that they
had previously consumed more than 2 units of alcohol or suffered from (uncorrected) poor vision. Furthermore, they were to be excluded if they exhibited physiological abnormalities, refused to comply with the instructions given by the experimenter, or due to technical problems. Based on these criteria, 6 participants were not included in the analysis.

The remaining 60 participants were assigned a condition (Stimulation Type) according to their ID number (modulo 3). Since most of the exclusion cases mentioned above became apparent during the study, it was possible to allocate new participants to specific conditions to balance the three groups as evenly as possible: Motor Condition - 19 participants, Physio Condition - 21, Passive Condition - 20.

Participants were between 18 and 64 years old (mean 23.5, sd 6.8), 38 were female and 8.3% were left handed. Independent of their gender or handedness, all participants experienced the same three hand representations in the place of their own right hand. On a scale from 1 (“beginner”) to 7 (“expert”) programming experience was low (median = 1, IQR = 2) and computer expertise was rated medium (median = 4, IQR = 2). In regard to previous VR experiences, 12/60 participants rated 1 (on a scale from 1 “never” to 7 “very often”) and only 23/60 gave ratings > 3 (median = 3, IQR = 2).

To ensure the fullest possible understanding of the material the information form, consent form and all questionnaires were available in English, Spanish, and Catalán. The pre-recorded instructions were also available in all three languages.

This study was performed with approval by the Comissió de Bioètica de la Universitat de Barcelona. All participants gave their written informed consent and were offered a compensation of 10 Euros for their time.

Procedure

The experiment lasted 40 to 60 minutes. On average 22 minutes were spent in virtual reality.

Upon arrival the participants read an information sheet and signed their informed consent. They were then asked to fill out a demographic information form and the BAQ [110]. Then three electrodes were attached to the participants’ torso to provide ECG measurements, which were used for the heartbeat visualization in the Physio Condition and for calculation of their IS. The participants were asked to take a seat at a desk and put on the HMD. The experimenter then instructed the participants to comfortably place their right hand on the table in resting position and reminded them not to move it unless explicitly told to do so. Their left hand was placed on a stool beside the table, close to the controller. Fig 10.8 shows a participant in the experimental setup. A set of wireless noise-canceling headphones provided the participants with
10.4. MATERIALS AND METHODS

Figure 10.8: **Experimental setup.** Participants were seated at a desk, with their right hand on the table top and their left hand on a controller.

pre-recorded instructions throughout the study.

When the experiment started the participants saw a virtual living room from first person perspective and appeared to be sitting at a table. The participants were asked to look around, familiarize themselves with the environment, and relax. During this time the experimenter verified that the electrodes were attached properly and the ECG data was collected correctly. Then the participants were instructed to perform a heartbeat counting task (mental tracking method explained in the Measurements section). Next the participants were asked to perform a first estimation of the position of their right hand to establish a baseline for the calculation of proprioceptive drift. Once these baseline measurements were established the hand trials began.

The participants were asked to look down towards their own right hand on the desk, where they could see a virtual model of a right hand. In the Motor Condition the participants were given instructions on how to move their hand, look at it from different angles, wriggle their fingers and observe how the virtual hand reacted. Finally, they were asked to place their hand back on the desk in resting position. In both the Passive Condition and the Physio Condition hand tracking was deactivated and the participants were reminded to keep their own hand motionless on the desk. In VR the unmoving virtual hand was shown in similar position and pose as their own hand. The participants were instructed to focus on the virtual hand, and asked to imagine how it would respond if they lifted and moved their hand.
CHAPTER 10. PHYSIOMOTOR PUBLICATION

After experiencing the virtual hand for 1 minute the participants were asked to turn their head and look straight forward and the virtual hand was removed. Then they were asked to look back down and perform another position estimation of their right hand. Thereafter, participants rated the items of the interactive questionnaire (Q1-Q4 in Table 10.1) by rotating the controller with their left hand. A new hand then appeared on the table and the above process was repeated.

The participants completed 5 hand trials, rotating through 3 different Visual Appearances (hand representations A, B and C in Fig 10.6), with the degree of realism either first decreasing then increasing (A-B-C-B-A), or vice-versa (C-B-A-B-C). Hereby the first two trials served as practice-trials and allowed participants to get acquainted with the two hand representations that they would experience last (to allow direct comparisons). In trials 3-5 they could then rate these hand representations accordingly. Hence, only data from trials 3 to 5 was included in the analysis.

After all trials were concluded the participants responded to the remaining questions in our questionnaire (Q5-Q10 in Table 10.1), were debriefed, and received a remuneration.

Results

Due to the mixed study design the data was analyzed with a mixed-measures ANOVA. Effects and interactions are assumed to be significant at $p < 0.05$.

In summary, our results show that the ownership illusion is indeed reduced by unrealistic-looking virtual hand representations (H1). While there was no significant effect of Stimulation Type, our data suggests that visual-motor synchrony (Motor Condition) may to some degree compensate for lack of connectivity of a virtual hand representation (H2). However, this effect is not significant and much weaker than expected. Finally, while the cardio-visual feedback (Physio Condition) was noticed, it did not show any impact on the VHI (H3). More details on these findings are given below.

Perceived Ownership

Ownership was evaluated mainly based on ratings of Q1 (“I felt as if the virtual hand might be part of my body.”). Responses to this statement revealed a significant main effect of Visual Appearance (Hands: A, B, C) (ANOVA: $F(1.62, 92.18) = 58.35, p < 0.05$). Paired t-tests revealed that all visual representations were experienced as significantly different to each other ($p < 0.01$). Corrections for multiple comparisons were made using the Holm method. The highest body ownership scores were given for the most realistic hand (A), while the lowest ones were reported for the disconnected hand (C). These ratings
10.4. MATERIALS AND METHODS

Figure 10.9: Ownership ratings (Q1) by Visual Appearance (left) and Stimulation Type (right). Q1 ("I felt as if the virtual hand might be part of my body.") is regarded as the main item in the ownership questionnaire. Left: There was a significant effect of Visual Appearance on perceived ownership. Across all conditions participants reported ownership of the most realistic hand (A). The abstract hand (B) was rated significantly lower, and the disconnected hand (C) mostly inhibited the occurrence of the ownership illusion. Right: There was no significant effect of Stimulation Type. Averaging responses for all hand representations results in predominantly positive ownership ratings in Motor Condition and at least 50% positive ratings in Physio Condition and Passive Condition. No significant interaction between Visual Appearance and Stimulation Type could be found. Ratings were given on a 7-point Likert scale from -3 (strongly disagree) to 3 (strongly agree).

are shown in Fig 10.9 (left) with the following median values: Hand A = 2.5, Hand B = 1.0, Hand C = -1.0).

There was no significant effect of Stimulation Type (Conditions: Motor, Physio, Passive; shown in Fig 10.9, right), neither did we find any significant interaction between Visual Appearance and Stimulation Type. However, we would like to point out a (non-significant) difference between conditions, which can be seen in Fig 10.10: while ownership significantly decreased with ownership, it appears that for Motor Condition 50% of participants still indicated positive degrees of ownership (>= 0) even for the disconnected hand (C). From this we could interpret that visual-motor synchrony may to some degree compensate for the lack of connectivity.

Note that the phrasing of Q1 differs slightly from corresponding items in earlier questionnaires since it includes the word “might”. This could potentially lead to a slightly different interpretation. For instance, a participant could rate this positively thinking “this (realistic) hand looks like it might be part of my body, but it doesn’t feel like it”. However, findings from Q1 were corroborated
CHAPTER 10. PHYSIOMOTOR PUBLICATION

Figure 10.10: Ratings of Ownership (Q1) by Visual Appearance evaluated for each Stimulation Type. In all conditions (left: Motor Condition, center: Physio Condition, right: Passive Condition) the most realistic hand representation (A) received high ownership ratings. Ratings for hand B were significantly lower, but still predominantly positive in Motor Condition and Passive Condition and only slightly lower in Physio Condition. The disconnected hand C led to a high variance of responses in Motor Condition, with half of the participants indicating some degree of ownership. There was no evidence of ownership of hand C in the vision-only conditions.

by responses to Q2 (“I felt like the virtual hand was NOT part of my body.”), which can be considered a more strict and negated variant thereof. For this item there was again a significant main effect of Visual Appearance (ANOVA: $F(1.61, 91.88) = 37.37, p < 0.05$). Paired t-tests with corrections using the Holm method revealed significant differences between all hands ($p < 0.01$), with ownership being the highest for the realistic hand (hand A median: -2.0), followed by the abstract hand (hand B median 1.0), and lowest for the disconnected hand (hand C median: 2.0). Note that the scale was inverted. There was again no significant interaction or main effect of Stimulation Type.

The item Q3 (“I felt as if I had more than one right hand.”) is considered a control question. This question was predominantly rated with disagreement (median ratings between -2 and -3) and no significant effects were found, neither of Visual Appearance, nor of Stimulation Type.

In regard to the ownership ratings in Physio Condition, we wish to acknowledge a limitation of the cardio-visual feedback. The change of color was applied to the surface of the hand, so it may have been less noticeable on the hand representations of lower realism, which featured less connected areas and therefore a smaller overall surface. In particular on Hand C, which consisted only of finger tips and the stump of an arm, the effect of the feedback may have been severely limited. While participants were asked whether they had noticed the color change to ensure that the effect was visible (as discussed later), they were not asked whether they had noticed it for each hand representation. An alternative approach could have been to render a red outline
around the body, as was done by Aspell et al. [7]. A second option that could be explored in future is to inform participants about the feedback and the meaning thereof and observe whether this knowledge of seeing their heartbeat on the virtual hands changes their perception thereof.

**Perceived Agency**

The fourth item of the questionnaire in Table 10.1 measures the participant’s feeling of agency over the virtual hand: Q4 “I felt like I might be able to control the virtual hand as if it was part of my own body.”. Here we found a significant main effect of both Visual Appearance (ANOVA: $F(1,64,93.20) = 33.84, p < 0.01$), as shown in the left part of Fig 10.11, and Stimulation Type (ANOVA: $F(2,57) = 3.39, p < 0.05$), as shown on the right. Paired t-tests with Holm correction reveal all hands to be significantly different from each other ($p < 0.01$) (median ratings: A = 3.0, B = 2.0, C = 1.0), with stronger agency for more realistic hands. Unpaired t-tests with Holm correction revealed significantly higher agency in **Motor Condition** (median: 2.0) than in both vision-only conditions ($p < 0.01$). This is not surprising since in this condition the virtual hand moved in synchrony with the participant’s hand. There was no difference between **Physio Condition** (median: 1.0) and **Passive Condition** (median: 1.0). In these vision-only conditions the virtual hand was immobile and participants were instructed to keep their hands still. Hence, the feeling of agency stems from them imagining what would happen if they moved their hand.

There was a significant interaction between Visual Appearance and Stimulation Type (ANOVA: $F(3.27,93.20) = 3.50, p < 0.05$). This interaction effect shows that under congruent visual-motor stimulation (**Motor Condition**) participants reported a high degree of experienced agency, independent of the degree of realism of the virtual hand (Fig 10.12, left). However, in the **Physio Condition** (Fig 10.12, center) and **Passive Condition** (Fig 10.12, right), agency was higher for the realistic virtual hand (A) compared to the unrealistic ones (B, C).

**Proprioceptive Drift**

Drift was measured as the difference between the indicated position of the real hand after each trial, and a first position indication made before seeing any virtual hand (baseline). Negative values correspond to a drift towards the virtual hand. The virtual hand was located 10 cm to the left of the participants’ real hand (i.e., the virtual hand was closer to the participant’s mid-line).

We found no significant main effects and no significant interaction of Visual Appearance or Stimulation Type. Measurements are visualized in Fig 10.13.
Figure 10.11: Agency ratings (Q4) by Visual Appearance (left) and Stimulation Type (right). Left: Perceived agency was significantly affected by the realism of the virtual hand, with higher realism eliciting a stronger feeling of agency. Right: There was evidence of agency in all condition. Unsurprisingly the Motor Condition achieved highest scores. The other two conditions elicited lesser degrees of agency. Agreement to item Q4 was indicated on a 7-point Likert scale.

with drift by Visual Appearance on the left (mean drift: \( A = -0.06, B = -0.06, C = -0.06 \)) and Stimulation Type on the right (mean drift: Motor Condition = -0.07, Physio Condition = -0.05, Passive Condition = -0.06). Values are given in meters, so for instance -0.06 corresponds to a 6cm-drift towards the virtual hand.

Interestingly, participants consistently indicated their real hand to be located about midway between virtual hand their actual hand position. This was independent of Stimulation Type and Visual Appearance, and there was no correlation between ownership ratings and proprioceptive drift. This result may be interpreted in line with the finding of Rohde et al. [104], namely that proprioceptive drift can occur in absence of ownership. Furthermore, it was reported by Perez-Marcos et al. [95] that proprioceptive drift was more related to the distance between the real and virtual arm, than to the illusion of ownership. From our experiment it might appears that focusing on a virtual hand that is located in a plausible position in relation to the own body, but with a slight offset from the own hand, was enough to recalibrate the participants’ body image even without perceiving the presented hand as their own.

However, we think it more likely that the proprioceptive drift is a carry-over effect. Between trials the virtual hand was hidden from view and the participant was asked to look around to move his focus away from the location
10.4. MATERIALS AND METHODS

Figure 10.12: Agency ratings (Q4) by Visual Appearance per Stimulation Type (Motor Condition, Physio Condition, Passive Condition). Across all conditions agency ratings were high for the most realistic hand representation, even when no hand tracking was supported. A significant interaction between Stimulation Type and Visual Appearance indicates that congruent visual-motor stimulation consistently led to strong agency, independent of the virtual hand representation. However, in vision-only conditions (Physio Condition, Passive Condition) agency significantly decreased with realism.

of the virtual hand. However, without being able to look at their own hands in between, this may not have been enough to trigger a proprioceptive update and remove the drift caused by a previous hand representation.

Interoceptive Sensitivity and Body Awareness

Awareness of processes in the own body was quantitatively evaluated based on the measure for IS [129]. This is a number between 0 and 1, where a higher value means more similarity between counted and recorded heartbeats and thus higher sensitivity to internal processes. Most participants counted less heartbeats than were measured (only about half as many), with an average IS of 0.65. Overall, participants found the counting task rather hard with median ratings to Q8 (“It was easy for me to count my heartbeats.”) between 2 (Motor Condition, Physio Condition) and 3 (Passive Condition) on a scale from 1 (“strongly disagree”) to 7 (“strongly agree”). The majority assessed their own counting performance as mediocre: the median rating to Q9 (“I did well in the heartbeat counting task.”) was 4.0 (Likert scale from 1-7). Self-assessed performance and IS are weakly correlated (Pearson’s r(178) = 0.16, p < 0.05).

The BAQ [110] served as a qualitative measure for body awareness. We summed the average of all 18 questionnaire items and calculated a ratio. This again results in a number between 0 and 1, with a higher value indicating higher body awareness. Based on this measure mean body awareness was 0.60.
Figure 10.13: **Proprioceptive Drift by Visual Appearance (left) and Stimulation Type (right).** Proprioceptive drift was measured as the error between the indicated position of the participant’s real right hand after experiencing the virtual hand representations, compared to a baseline measurement. A smaller value corresponds to a position estimation that is closer to the virtual hand. Measurements were remarkably similar for all Visual Appearances (left) and Stimulation Types (right), suggesting a constant proprioceptive drift of 6 cm towards the virtual hand.

While the mean values are similar, we found no correlation between IS and BAQ (Pearson’s $r(178) = 0.11, p = 0.15$). Furthermore, there was no correlation of either IS or BAQ with Q1, or proprioceptive drift. Hence, we could not reproduce the earlier finding that people with higher IS are less likely to experience ownership illusions or experience them less strongly [2, 129].

However, we wish to mention that some participants admitted to feeling their pulse on their face due to the tight straps of the HMD. Similarly, some participants may have felt their pulse in the back of their leg when sitting on the chair. This could have led to an artificially high IS that may not align with the participant’s susceptibility to the ownership illusion, and could also explain the lack of correlation between IS and BAQ. However, if we assume that IS has an effect on the ownership illusion, and IS correlates with BAQ, it follows that we should also find a connection between BAQ and ownership ratings. Unfortunately, we found no such evidence.

**Perception of Cardio-Visual Feedback**

To ensure that the visualization of the heartbeat was adequate, participants were asked to rate a series of questions at the end of the experiment. Responses to Q5 (“There were times when I noticed that my virtual hand was changing color/flashing.”) show a significant main effect of Stimulation Type (Kruskal-
10.4. MATERIALS AND METHODS

Figure 10.14: Perception of cardio-visual feedback (left) and recognition of this feedback as own heartbeat (right). Responses to the question “There were times when I noticed that my virtual hand was changing color/flashing.” (left) indicate that the cardio-visual feedback in Physio Condition was indeed noticed. However, the disagreement towards “It seemed as if the flashing of the virtual hand was my heartbeat.” (right) shows that this color-change was not associated with the own heartbeat.

Wallis: $X^2(2) = 1.71, p < 0.01)$. A post-hoc Wilcoxon test with Bonferroni correction showed differences between all conditions ($p < 0.01$). This question was rated on a scale from 1 (“strongly disagree”) to 7 (“strongly agree”). With a median of 6.0 most participants correctly reported the color change in Physio Condition, where cardio-visual feedback was actually provided (see Fig 10.14, left). In both other conditions no such flashing was reported (Motor Condition median = 2.0, Passive Condition median = 1.0). As was already mentioned in context of the ownership questionnaire ratings, the fact that these questions were not asked in respect to individual virtual hand representations means that we cannot know whether the cardio-visual feedback was effective for all Visual Appearances.

In response to the Q6 (“It seemed as if the flashing of the virtual hand was my heartbeat.”) we also found a significant main effect of Condition on Flashing was Heartbeat (Kruskal-Wallis: $X^2(2) = 11.39, p < 0.01$). Here a post-hoc Wilcoxon test with Bonferroni correction showed significantly higher ratings ($p < 0.01$) in Physio Condition (median = 2.0) than in both other conditions (Motor Condition median = 1.0, Passive Condition median = 1.0). However, on a scale from 1 to 7 those ratings still indicate disagreement (Fig 10.14, right), meaning that even when the color change was noticed, participants did not recognize this as their own heartbeat.

157
CHAPTER 10. PHYSIOMOTOR PUBLICATION

Preferred hand representation

As final part of the experiment, participants were asked to indicate their preference for each of the virtual hand representations in pairwise comparisons. The vast majority of participants preferred the most realistic hand representation (A) over the other two (hands B and C). They explained that it looked more real and similar to the own hand, in particular due to details such as fingernails. Out of 60 participants, 7 preferred the hand of medium realism (B) to hand A. Reasons given include that the hand pose of B seemed more like their own, they felt more in control of hand B (in Motor Condition). Furthermore, comments were made that suggest that for some the realistic hand (A) may be in the uncanny valley while the abstract hand (B) was repeatedly described as a wooden hand. Hand C was mostly perceived as unrealistic and even scary. Nevertheless 2 participants preferred this over hand A, arguing that C was to easier control and provided a more interesting experience. Most participants also preferred hand B over hand C, with 5 exceptions who felt more in control and felt capable of more precise movements when seeing only the pointy fingertips. Notably most participants with this opinion (4/5) experienced hand C in the Motor Condition.

10.5 Conclusion

In a study with 60 participants we found that Visual Appearance has a strong impact on the ownership illusion, whereas the Stimulation Type does not. A realistic-looking virtual hand (hand A) received high ratings for perceived ownership and agency, independently of the type of stimulus (visual-motor, visual-proprioceptive, or cardio-visual). Unrealistic visual features, such as abstract, blocky fingers or missing connectivity of the hand, led to a significant decrease in ownership ratings. Agency ratings were similarly affected in the vision-only conditions (Physio Condition, Passive Condition). However, agency remained high for unrealistic hands in Motor Condition. Ownership ratings further indicate that visual-motor correlation (Motor Condition) may to some degree compensate for an unrealistic hand representation with missing connectivity. Furthermore, we found no evidence of a positive effect of the heartbeat visualization (Physio Condition) on the ownership illusion. The benefit of visualizing interoceptive signals may therefore be limited to embodiment of body-external and non-corporeal objects.
Bibliography


BIBLIOGRAPHY


161
BIBLIOGRAPHY


[33] Francesca Ferri, Antonio Maria Chiarelli, Arcangelo Merla, Vittorio Gallese, and Marcello Costantini. The body beyond the body: expectation of a sensory event is enough to induce ownership over a fake hand. Proceedings of the Royal Society of London B: Biological Sciences, 280(1765):20131140, 2013. 16


[36] Tiare Feuchtner, Jörg Müller, and Mel slater. How different multisensory stimuli and visual realism modulate the virtual hand illusion. In submission to Transactions on Computer-Human Interaction (TOCHI). ACM, 2018. vii, ix, 8, 9, 26, 29, 47, 54, 72, 74, 77, 81, 83


162


BIBLIOGRAPHY


BIBLIOGRAPHY


166


[96] Valeria I Petkova and H Henrik Ehrsson. If i were you: perceptual illusion of body swapping. PloS one, 3(12):e3832, 2008. 14, 18, 31, 75, 82, 83, 93, 94, 115, 137


170


BIBLIOGRAPHY