Embodiment Sensitivity to Movement Distortion and Perspective Taking in Virtual Reality

THÈSE N^{O} 7180 (2017)

PRÉSENTÉE LE 17 FÉVRIER 2017

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS
GROUPE SCI IC RB
PROGRAMME DOCTORAL EN ROBOTIQUE, CONTRÔLE ET SYSTÈMES INTELLIGENTS

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Henrique GALVAN DEBARBA

acceptée sur proposition du jury:

Prof. A. Ijspeert, président du jury
Dr R. Boulic, Dr B. Herbelin, directeurs de thèse
Dr A. Lécuyer, rapporteur
Prof. F. Tecchia, rapporteur
Dr P. Pu Faltings, rapporteuse

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2017

Acknowledgements

First and foremost, I am grateful for the time and effort that my thesis directors, Dr. Ronan Boulic and Dr. Bruno Herbelin, invested on me. Our discussions were always entertaining and motivating. I am also grateful to Dr. Anatole Lécuyer, Prof. Franco Tecchia, Dr. Pearl Pu and Prof. Auke Ijspeert for accepting the invitation to review this document and taking part in my thesis defense.

My thanks to the colleagues in the Immersive Interaction Group: Eray Molla, Nan Wang, Junghyun Ahn, Joan Llobera, Pawan Harish, Mentar Mahmudi, Benoît Le Callennec and Claire-Lise Coscia; the EDRS secretary Corinne Lebet, and to all colleagues in the Laboratory of Cognitive Neuroscience and Campus Biotech, in particular to Roy Salomon, Andrea Serino, Javier Bello Ruiz, Petr Grivaz and Lucie Bréchet, who had direct influence in the work presented here.

I would also like to thank the Master and Bachelor students whom I had the privilege to advise in the past 4 years: Sidney Bovet, Sami Perrin, Jad-Nicolas Khoury, Florian Junker, Samuel Gruner, Yunpeng Zhou and Fabien Schmitt. In particular, I would like to acknowledge the equal contribution of Sidney Bovet in the experiment presented in Section 4.3, Sami Perrin in the experiment presented in Section 3.1, and Jad-Nicolas Khoury and Sami Perrin in the experiment presented in Appendix B.

I wish to express my sincere thanks to Sandra Roux and her family, who opened their home to me when I needed the most, and nurtured a prolific friendship during the past 4 years. I would also like to praise Sandra on her role in the registrar's office of EPFL.

Last but not least, I thank my parents, Ademir and Inês Debarba, and my wife, Monique Sonego, for their love, encouragement and support.

Lausanne, 11 October 2016 Henrique Galvan Debarba

Abstract

Despite recent technological improvements of immersive technologies, Virtual Reality suffers from severe intrinsic limitations, in particular the immateriality of the visible 3D environment. Typically, any simulation and manipulation in a cluttered environment would ideally require providing feedback of collisions to every body parts (arms, legs, trunk, etc.) and not only to the hands as has been originally explored with haptic feedback. This thesis addresses these limitations by relying on a cross modal perception and cognitive approach instead of haptic or force feedback. We base our design on scientific knowledge of bodily self-consciousness and embodiment. It is known that the instantaneous experience of embodiment emerges from the coherent multisensory integration of bodily signals taking place in the brain, and that altering this mechanism can temporarily change how one perceives properties of their own body. This mechanism is at stake during a VR simulation, and this thesis explores the new venues of interaction design based on these fundamental scientific findings about the embodied self. In particular, we explore the use of third person perspective (3PP) instead of permanently offering the traditional first person perspective (1PP), and we manipulate the user-avatar motor mapping to achieve a broader range of interactions while maintaining embodiment. We are guided by two principles, to explore the extent to which we can enhance VR interaction through the manipulation of bodily aspects, and to identify the extent to which a given manipulation affects the embodiment of a virtual body.

Our results provide new evidence supporting strong embodiment of a virtual body even when viewed from 3PP, and in particular that voluntarily alternating point of view between 1PP and 3PP is not detrimental to the experience of ownership over the virtual body. Moreover, detailed analysis of movement quality show highly similar reaching behavior in both perspective conditions, and only obvious advantages or disadvantages of each perspective depending on the situation (e.g. occlusion of target by the body in 3PP, limited field of view in 1PP). We also show that subjects are insensitive to visuo-proprioceptive movement distortions when the nature of the distortion was not made explicit, and that subjects are biased toward self-attributing distorted movements that make the task easier.

Key words: virtual reality, embodiment, sense of agency, body ownership, first and third person perspectives, human-computer interaction, visuomotor contingencies, virtual body.

Résumé

Malgré les récents progrès des technologies d'immersion, la réalité virtuelle souffre de limitations intrinsèques liées à l'immatérialité de l'espace visuel en 3D. Typiquement, la simulation et l'interaction dans un espace virtuel encombré devrait idéalement fournir un retour tactile pour les collisions des objets avec toutes les parties du corps, et pas seulement sur les mains comme généralement exploré avec des interfaces haptiques. Cette thèse vise à réduire ces limitations en adoptant une approche cognitive et perceptive, au lieu de développer un système de retour de force ou tactile. La recherche en sciences cognitives sur la conscience de soi et de son corps a montré que l'expérience subjective d'incarner un corps est issue de la cohérence des signaux corporels multimodaux qui sont intégrés dans le cerveau pour construire la représentation de soi, et que l'altération de ce mécanisme peut modifier la perception du corps. C'est ce même mécanisme qui est en jeu lors d'une simulation en réalité virtuelle, et nous proposons d'appliquer ces connaissances au design d'interaction en environnement virtuel. En particulier, nous explorons l'utilisation de la perspective à la troisième personne (3PP) en complément de la traditionnelle vue en première personne (1PP), et nous manipulons la concordance motrice entre l'utilisateur et son avatar pour étendre les possibilités d'interaction tout en maintenant l'incarnation dans le corps virtuel. Nous somme guidés par deux principes ; améliorer l'interaction en RV en exploitant la manipulation de conscience de soi, et identifier l'influence de ces manipulations sur la perception d'incarner un avatar.

Nos résultats apportent des éléments nouveaux soutenant l'idée qu'un fort sentiment d'incarner un corps virtuel est possible même s'il est vue en troisième personne, et ce en particulier quand l'utilisateur peut volontairement contrôler le passage de 1ere à 3eme personne. L'analyse détaillée des mouvements d'interaction montre un comportement similaire dans les deux conditions de perspective, avec uniquement des qualités et inconvénients spécifiques à chacune (p.ex. occlusion de la cible par le corps en 3PP, champ de vision réduit en 1PP). Nous montrons aussi que les sujets sont insensibles aux distorsions visuo-proprioceptives du mouvement quand la nature de celle-ci n'est pas évidente, et que les sujets considèrent facilement que le mouvement déformé est le leur si la distorsion rend la tâche plus facile.

Mots clefs : Réalite virtuelle, incarnation d'un corps, sens de l'agentivité, perspectives à la première et la troisième personne, interaction homme-machine, contingences visuo-motrices, corps virtuel.

Contents

Acknowledgements **i**

Abstract (English/Français) **iii**

List of figures **xi**

List of tables **xiii**

1 Introduction **1**
1.1 Research problem and approach 2
1.2 Scope of the Thesis 3
1.3 Contributions 4
1.4 Organization 4

2 Background and Literature Review **7**
2.1 Virtual Reality and Human Perception 7
2.1.1 Immersion 7
2.1.2 Presence 8
2.1.3 Human Perception applied to VR 9
2.2 Embodiment and self-consciousness 10
2.2.1 Defining the self 10
2.2.2 Sense of Embodiment 11
2.2.3 Manipulating the Bodily Self 12
2.2.4 Perpective taking 13
2.3 The self as an actor 15
2.3.1 Internal Models of movement control 15
2.3.2 Comparator model and the sense of agency 16
2.3.3 Sensorimotor adaptation 17

3 Increasing the Awareness of the Virtual Environment **19**
3.1 Non-planar Projection with HMDs 19
3.1.1 Related Work 21
3.1.2 Materials and Methods 22
3.1.3 Results 23
3.1.4 Discussion 26

3.2 Combining 1PP and 3PP 27
3.2.1 Related Work 28
3.2.2 Design Space for Combining Perspectives 28
3.2.3 Design 1: Simultaneous Perspective Interface 30
3.2.4 Design 2: Alternating Perspective Interface 32
3.2.5 Discussion 33
3.3 Synthesis and conclusion 34

4 Perspective Taking and Embodiment **37**
4.1 Related Work 38
4.2 Experiment 1: Perspective and Visuo-motor Synchrony 38
4.2.1 Materials and Methods 39
4.2.2 Results 45
4.2.3 Discussion and Conclusion 46
4.3 Experiment 2: Perspective and Multimodal Congruence 49
4.3.1 Materials and Methods 49
4.3.2 Results 56
4.3.3 Discussion 60
4.4 Synthesis and Conclusion 64

5 Movement Distortion and Embodiment **69**
5.1 Related work 70
5.2 Distortion Model 71
5.3 Materials and Methods 73
5.4 Experiment 1: Just Noticeable Difference 77
5.5 Experiment 2: Task Difficulty 79
5.6 Results 80
5.7 Discussion 82
5.8 Conclusions 83

6 Conclusion **85**
6.1 Contributions 86
6.2 Outlook 87
6.3 Limitations and Future Work 88

A Appendix **91**
A.1 Neural signatures of self-movement and movements distortion in embodied VR 91

B Appendix **93**
B.1 Manipulating movement precision 93
B.1.1 Distortion Model 93
B.1.2 Materials and Methods 95
B.1.3 Results and Conclusions 96

C Appendix **99**

D Appendix **103**

E Appendix **107**

Bibliography **111**

List of acronyms **123**

Curriculum Vitae **125**

List of Figures

1.1 Scope of the thesis. 3

2.1 Human-Virtual Environment interaction loop. 8
2.2 Illustration of the rubber hand illusion. 13
2.3 Full body ownership illusion. 14
2.4 Full body ownership illusion in 3PP. 15
2.5 Internal model of action 16

3.1 Virtual environment and screenshots of tested projections. 20
3.2 Overview of the system setup. 23
3.3 Non-planar projections performance results. 24
3.4 Non-planar projections collision results 25
3.5 Non-planar projections embodiment and simulation sickness results 25
3.6 Design space for combining 1PP and 3PP viewpoints 29
3.7 Possible placement of viewports 31
3.8 Proof of concept of our simultaneous perspective interface. 32
3.9 Screenshots of perspective alternation 33

4.1 3PP camera behavior. 40
4.2 Overview of perspective conditions and setup. 41
4.3 Criteria defining the two stages of a reaching. 43
4.4 Boxplots of questionnaire results 45
4.5 Mean difference for Mean Time response. 46
4.6 Mean difference for Mean Time Stage 1 and 2 responses 47
4.7 End effector of preference for each target per perspective 48
4.8 Experiment setup and scene overview. 50
4.9 Perspective conditions. 51
4.10 Overview of the session stages. 54
4.11 Reported senses of agency, ownership, self-location and threat 57
4.12 Sense of *self-location* at different levels of perspective and perspective order. . . 59
4.13 Floor fall event locked GSR variation. 60
4.14 GSR difference resulting from the floor fall event. 61
4.15 Example of a GSR record of a session. 61
4.16 Boxplot of the Mental Ball Drop results. 62

4.17 Boxplot of the time to reach floor and air targets per perspective. 63
4.18 Proportion of time spent in 1PP for each stage of the Alternate session. 64

5.1 Overview of the reaching distortion function. 72
5.2 Overview of the reaching experiment setup. 74
5.3 Overview of the reaching experiment trial. 75
5.4 Samples of adaptive staircases. 78
5.5 Bar plots of the JND thresholds in different scales. 81
5.6 Subjective evaluation of the difference in difficulty due to movement distortion. 81

B.1 Overview of the pointing experiment trial. 94
B.2 Overview of the pointing distortion function. 94
B.3 Overview of the pointing experiment setup 96
B.4 Pointing experiment JND and PSE results. 97
B.5 Pointing experiment relation of self-attribution and perception of errors. 98

List of Tables

3.1 Classification according to the proposed design space. 34

4.1 Questionnaire and results 42
4.2 Embodiment questionnaire applied in the end of each session 55
4.3 Results per experimental condition. 57
4.4 Post-experiment responses for the VMT group. 65
4.5 Perspective taking and embodiment statistical significance tests summary. . . . 67

5.1 Estimated points of Just Noticeable Difference (JND) for different scales. 80
5.2 Estimated points of Just Noticeable Difference (JND) and Point of Subjective Equality (PSE) for distance proportion difference scale. 82
5.3 Potential bias according to distortion orientation and movement direction . . . 83

1 Introduction

We spontaneously experience our body as a consistent and seemingly immutable representation of ourselves in space. We do not expect our body to change in shape or size, yet it does. We all undergo considerable body reshaping while growing from infancy into adulthood, but to the extent we can record, our body representation felt solid at any point of this transformation.

The instantaneous experience of embodiment emerges from the coherent multisensorial integration taking place in the brain, which has been referred to as bodily self-consciousness (the pre-reflective sensation of being the subject of an experience) [Legrand, 2006, Blanke, 2012, Blanke et al., 2015]. Experimental protocols have shown that this body representation is much more malleable than commonly assumed. For instance, conflicting sensorial stimulation can temporarily change how one perceives properties of their own body (i.e. an altered bodily self-consciousness). Notably, it can lead to the illusion of owning a fake – either material or virtual – limb [Botvinick and Cohen, 1998], body [Slater et al., 2010b], and even another individuals' body [Petkova and Ehrsson, 2008].

In the rubber hand illusion [Botvinick and Cohen, 1998], the synchronous stroking of a (visible) rubber hand and the (occluded) real hand provides visuo-tactile congruence to the subject, while causing a visuo-proprioceptive conflict. That is, relative to the seen rubber hand, the subject feels the touch in a congruent body region and time, but the global location of the hand does not match. Eventually, the subject feels ownership over the fake limb, which is accompanied by the feeling that the real hand is now located closer to the rubber hand. The same happens if visuo-motor or visuo-proprioceptive congruence (active or passive movement of the hand) is used in lieu of the visuo-tactile stimulation.

Virtual reality (VR) is especially competent in producing these bodily illusions. For instance, compared to physical reality, the development of a VR application supporting the control of a virtual limb or a virtual body, which yields visual, motor and proprioceptive congruence, is relatively straightforward and much more malleable. By using current motion capture equipment and animation algorithms, it is possible to give full control of a virtual (non) humanoid body to a subject, effectively augmenting the means through which one can deceive the multisensorial mechanisms that give rise to the embodied self.

Nonetheless, there is a longstanding interest of VR for what composes the so called feeling of "presence", which is commonly defined as the feeling of "non mediation" [Lombard and Ditton, 1997] and of "being there", in the virtual world. To that extent, the sense of embodying a virtual body addresses a complementary subject, that of a self-body relation, while presence approaches the mental state of being located in the VE, i.e. the self-environment or body-environment relation [Kilteni et al., 2012a]. Evidence suggests that these are closely related, and that having and controlling a virtual body in a VR simulation is among the main factors driving the experience of presence [Slater et al., 2010a].

As a matter of fact, early conceptualization of presence have emphasized the role of the body. In the "Cyborg's Dilemma", [Biocca, 1997] approaches the concept of "self-presence", defined as the effect of the virtual environment on the perception of one's body, physiological states, emotional states, perceived traits and identity, i.e. the notion of a bodily self that can be modeled after a synthetic experience. Indeed, recent evidence shows that a congruent multisensory experience in VR can alter how one perceives her/his body shape [Kilteni et al., 2012b, Normand et al., 2011].

Adding to this discussion, this thesis explores how sensory and sensorimotor discrepancies affect the embodied self in VR, and how this knowledge could be applied to the design of embodied VR interaction.

1.1 Research problem and approach

What is at the core of an effective VR experience is the idea that people forget about the technological mediation and experience their visit as "being there" in a tangible world. Simulations are usually performed with a first person perspective view into the 3D space, which is supposed to be "natural" as the user perceives and moves in the virtual world as in reality.

But VR suffers from severe limitations, in particular the immateriality of the visible 3D environment. For example, haptic displays may be competent at providing touch and/or force feedback to probes and even hands (such as in surgical applications), but are still unable to provide realistic full-body haptic feedback as tactile and bodily sensations are far wider in terms of skin surface. However, typical simulation and manipulation in a cluttered environment requires providing feedback of collisions to every body parts (arms, legs, trunk, etc.) and not only to the hands. This corresponds to a large family of applications, including virtual prototyping and training in complex environments (plane assembly, industrial system maintenance, aftermath intervention) which are still impaired by this problem.

In this thesis we seek to address these limitations by relying on a crossmodal perception approach. Instead of extending the research in the direction of haptic and force feedback, we propose to fork towards a complementary direction by exploring a complete change of viewing perspective. First, we favor a third person perspective (3PP) viewpoint instead of permanently offering the traditional first person perspective (1PP) of VR interactions. Second we act on the user-avatar mapping to achieve a broader range of interactions, effectively modifying the motor mapping of reaching tasks, which can be used to guide the user interaction with the virtual environment (e.g. prevent interpenetration of virtual body with virtual scene from happening). Our hypothesis is that such alternative approach can provide useful feedback on body posture and interaction with the virtual world during embodied immersive interactions, especially when involved in potentially complex virtual environments.

1.2 Scope of the Thesis

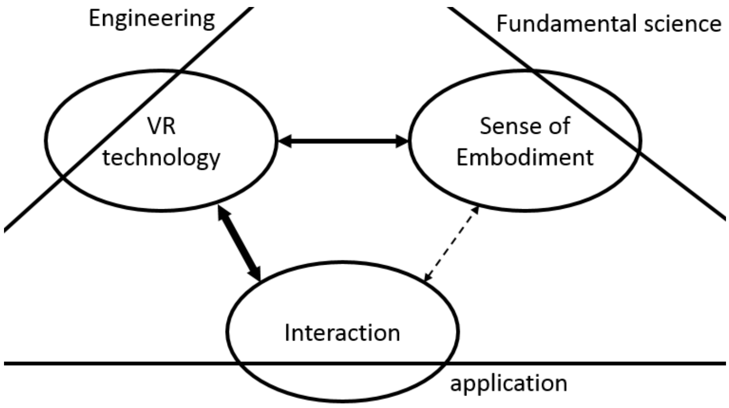


Figure 1.1 – The scope of this thesis is on how current knowledge on the sense of embodiment and VR interaction paradigms can impact each other.

We focus on how current knowledge on the sense of embodiment and VR interaction paradigms can impact each other (Figure 1.1). Specifically, we explore the new venues of interaction design based on fundamental scientific findings about the embodied self. In contrast, we do not create new VR technology nor do we explore embodiment at a neuronal/brain structure level. Moreover, although we confront our results with current ideas on embodiment and motor control, we do not put forward new theories on these topics.

We believe that the knowledge on how the brain represents the body is not only impactful to the fundamental research on cognitive neuroscience and to the field of neuroprosthetics, but it also has the potential to immediately impact the design of the fast growing market of VR applications. We argue that understanding the limits of embodiment and how it can be manipulated can lead to new venues for effective virtual reality interaction. Therefore, while most of current research uses VR to expand knowledge on the mechanisms of bodily self-consciousness, we also seek to bring this knowledge to the practical grounds of VR interaction. We are guided by two principles, to explore the extent with which we can enhance VR interaction through the manipulation of bodily aspects, and to identify the extent with which a given manipulation affects the embodiment of a virtual body.

1.3 Contributions

In this thesis we study the sense of embodiment and VR interaction when the "natural" mapping of oneself to a surrogate body is disrupted. In the interaction design side we explore visual feedback to convey structural information, and deviate movement to artificially control the virtual body relation with the environment. In the embodiment side we try to ensure that our interaction design decisions are compatible with the illusion of an altered bodily self-consciousness.

In particular, we provide new evidence supporting the embodiment of a virtual body controlled from 3PP. We hypothesize that this may be specifically related to the dynamics of sensorimotor contingencies relating the full real and virtual bodies, and to actively interacting with the virtual environment. The latter argument is based on the account of agency and embodiment proposed by [Synofzik et al., 2008b], in which the authors suggest that the conscious perception of ownership and agency is influenced by intentions. Moreover, our studies are the only ones to explore the role of full body control to the sense of embodiment of a virtual body seen from a 3PP. We also argue in favor of the possibility of dynamically alternating the point of view between 1PP and 3PP during the simulation. This approach intends to sum up the strengths of 1PP and 3PP in a single user interface.

In a second part of our research we investigate how one perceives self-generated movements when their visual feedback is altered using embodied VR, that is, when the virtual body does not perform the exact same movement as the person controlling it. Based on experiments, we propose distortion limits that are likely to be accepted by the subject as congruent to their actions. We also place our experiments in the context of current theories leveraging the role of movement monitoring, intentions, and retrospective inference to agency and self-attribution.

1.4 Organization

The remaining of this thesis is organized in 5 chapters:

Chapter 2 presents fundamental concepts that lay the base for this thesis. There we introduce: concepts of VR; how VR explores human perception for its own sake; current definitions of the self, of the sense of embodiment, and how it migth be affected by multisensory integration; we close by discussing motor awareness, motor control, and detailing motor aspects of embodiment.

In Chapter 3 we examine how the manipulation of visual feedback could be used in order to improve subjects awareness of a controlled virtual body posture and its relation with the virtual environment. In particular, we evaluate the use of non-planar projections as a means to increase the field of view, and how 3PP could be combined with 1PP to add up their advantages.

Chapter 4 explores how perspective (1PP/3PP) and visuo-motor congruency influence the sense of embodiment over a virtual body. We present two experiments on the subject. One of the experiments also include a perspective option that combines 1PP and 3PP, as proposed in Chapter 3. Additionally, we also expose performance differences and similarities between 1PP and 3PP.

In Chapter 5 we present the results of two experiments on self-attribution of hand movements with spatiotemporal distortions. In the experiments we manipulate the distance to the target by making it physically closer or farther, while the visual (apparent) distance is kept constant. The goal is to quantify people acceptance and subjective interpretation of these manipulations.

Finally, in chapter 6 we summarize and discuss the potential impact of this thesis. We finish by exposing ideas and intentions for the future of this research.

2 Background and Literature Review

In this Chapter we present concepts and theories on the subjects of virtual reality, self-consciousness and embodiment, and motor control. Our subject of research is multidisciplinar, and these topics provide the fundamentals for the understanding and the contextualization of this thesis.

2.1 Virtual Reality and Human Perception

The premise of virtual reality (VR) is to deliver a synthetic world that can be experienced as if it were real. Ideally, VR should mediate all input and output channels of a person to a point where she can no longer detect a discrepancy between the expected and rendered outcome to her actions. This i/o feedback loop is expected to register and interpret the user's actions and provide appropriate sensorial replacement (Figure 2.1), e.g. every time the user moves the head, a computer has to measure this motion and generate a new picture, coherent with the new viewpoint.

In this thesis we make extensive use of VR techniques to perform experiments on bodily perception. The development of VR is technically challenging. Mediating one's input to generate the corresponding output in VR incur on inevitable tracking and sensory rendering latency and imprecision. Below we discuss two base concepts of the field of VR, namely *immersion* and *presence*, as well as how VR relies on human perception to succeed.

2.1.1 Immersion

The concept of immersion refers to the objective level of sensory fidelity a VR system provides [Slater, 2003, Sanchez-Vives and Slater, 2005]. An extensive literature has studied, for instance, how the level of visual immersion depends on the system's rendering software and display technology. In that frame of mind [Pausch et al., 1993] compared head-mounted and stationary displays. [Bowman and McMahan, 2007] have chosen to study the level of visual immersion on application effectiveness by combining various immersion components such as field of

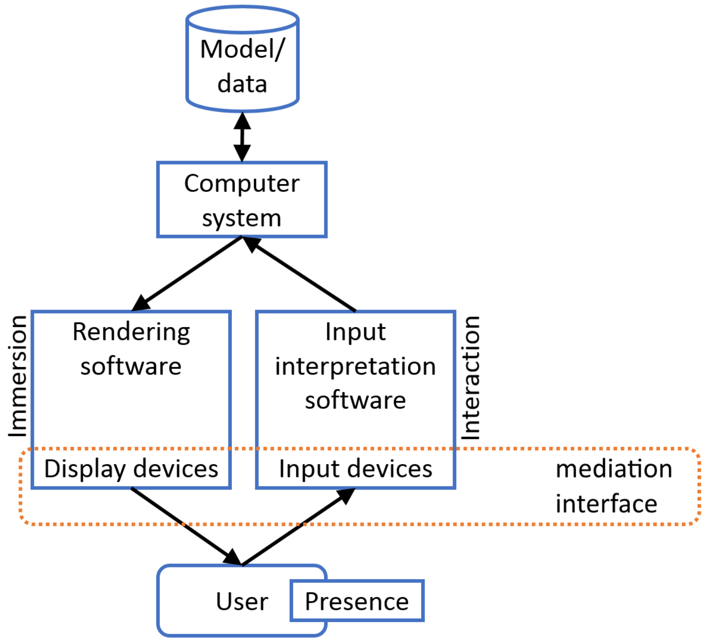


Figure 2.1 – Human-Virtual Environment interaction loop. Adapted from [Bowman and McMahan, 2007].

view, field of regard (total size of the visual field surrounding the user), display size, display resolution, stereoscopy, frame rate, etc.

2.1.2 Presence

The experience of a virtual environment (VE) through immersive aparatus can give rise to the sense of presence [Held and Durlach, 1992, Slater and Usoh, 1993], which has been described as the feeling of "being there", in the VE, and is marked by the "illusion of non-mediation" [Lombard and Ditton, 1997], when the equipment managing the feedback loop goes unnoticed in the user-VE relation. What is central to the state of presence is that the user *act in* and *react to* the VE as if it were real, despite the fact that the user *knows* that it is a simulation. For instance, the experiment proposed by [Meehan et al., 2002] expose subjects to a virtual pit, and the reactions that the VE triggers in the subject may be equivalent to those of a real exposure [Sanchez-Vives and Slater, 2005].

The concept of presence refers to a person's subjective psychological response to a VR system [Slater, 2003]. Therefore, albeit the fact that technical specifications of a VR system (how immersive it is) is a driving factor for the sense of presence, presence is above all the result of a person's psychological state, thus depending on many other factors. For instance, two users can have different experiences of presence with the same VR system, and the same user can have different experiences of presence in the same system at different times [Slater, 2003].

However, if the term "presence" is widely accepted as the label to design this feeling of "being there" [Heeter, 1992, Slater, 2003], its precise and specific understanding is still debated today. In this sense, [Slater, 2009] proposes two orthogonal components to presence, namely

- place illusion (PI) – encompasses the classical "being there" definition;
- plausibility (Psi) – the extent with which the participant feels that what happens is real.

[Slater et al., 2010a] performed an experiment to evaluate the relative importance of simulation elements to PI and Psi. They concluded that effective PI relates mainly to immersive apparatus and to having and controlling a virtual body (sensorimotor contingencies relating real and virtual body). Thus being tightly connected to the notion of an efficient feedback loop. In other words, sensorimotor feedback is the basic foundation for PI to occur; "the sense of 'being there' in a VE is grounded on the ability to 'do' there" [Sanchez-Vives and Slater, 2005]. On the other hand, effective Psi has been associated to illumination realism, with the virtual body still playing a prominent role. Hence, Psi seems to be especially related to higher order cognitive priors about what elements are contained in reality, and how one expect these elements to behave and look like. Both illusions can occur, together or independently, albeit participants' knowledge that the virtual environment is a simulation.

2.1.3 Human Perception applied to VR

Perception involves mechanisms that receive sensorial input and transform them from lower-level information (e.g. physical data) into higher-level information (e.g. shape and motion), and the processing of this information which is influenced by one's concepts, expectations and selective mechanisms (knowledge and attention) [Bernstein, 2013]. Thus, one's perception is influenced at different levels, from the most mechanical aspects of the sensorial apparatus, to one's current psychological state, to cognition. Notably, much of the information we experience as being collected from the external world may be product of brain inference [Ramachandran et al., 1991].

As a consequence, the VR feedback loop is effective even if its technical specifications are inferior to physical reality and physiological limits, i.e. to be effective, it only has to be as good as human perception and expectations. Ultimately, the subjective match of simulation and expectation can give rise to an "illusion of non-mediation" [Lombard and Ditton, 1997], which is at the core of the idea that one can feel to be present in the virtual world.

Perhaps, one of the most interesting samples of faulting perception is that of crossmodal illusions. For instance, in the well known ventriloquist effect, the synchronicity of the moving puppet mouth and the ventriloquist voice gives the perception that the sound is being projected from the puppet. This illusion implies that auditory perception can be shaped by vision, i.e. the perception of the ensemble predominates over a single modal input. VR researchers have long explored this sort of modal predominance to improve interaction. In a family of navigation techniques known as redirected walking, the predominance of visual over vestibular sensorial input is exploited in order to maximize the virtual space accessible through natural walking [Razzaque et al., 2001, Steinicke et al., 2010]. In a more recent approach, [Kohli et al., 2012] redirect the movement of an end effector to provide congruent visual and tactile stimulation, which seems to predominate over proprioceptive information.

Our approach to the problem of immateriality of VR draws from this literature. We rely on crossmodal perception, and explore a complete change of viewing perspective. First, we favor a third person perspective (3PP) instead of permanently offering the traditional first person perspective (1PP) of VR interactions. Second we act on the user-avatar mapping to achieve a broader range of interactions, effectively modifying the motor mapping of reaching tasks in order to guide the interaction with the virtual environment.

2.2 Embodiment and self-consciousness

Embodiment, as defined in the fields of philosophy of the mind and cognitive neuroscience, emphasizes the relevance of sensorimotor skills to the shaping of the mind and the subjective experience of having and controlling a body [Blanke and Metzinger, 2009]. To this extend, we walk through definitions of the self and of sense of embodiment; we present experimental protocols that manipulate the perception of the bodily self; finally, we give emphasis to the manipulation of point of view, which is especially relevant for Chapters 3 and 4.

2.2.1 Defining the self

In [Gallagher, 2000], Gallagher overviews two orthogonal components of the self, the minimal self and the narrative self. The former refers to the minimal necessary condition for the instantaneous experience of being a self, and is limited to what is accessible to immediate self-consciousness. The latter refers to higher level concepts defining a self-image, such as the auto-biographical views that persons build for themselves, thus having a narrative past and future.

The research we propose relates to the concept of a minimal self. More specifically, to the account of a minimal phenomenal self [Blanke and Metzinger, 2009], which is defined as "the experience of being a distinct, holistic entity capable of global self-control and attention, possessing a body and a location in space and time". Thus contemplating physical aspects that help to define a bodily self. Interestingly, this self distinction relates to the components of presence proposed by Slater [Slater, 2009]. Notably, the simulation mechanisms found to support place illusion (immersive apparatus and sensorimotor contingencies) overlaps with aspects that define the minimal phenomenal self [Blanke and Metzinger, 2009].

2.2.2 Sense of Embodiment

[De Vignemont, 2011] proposes the following general definition of embodiment: "E is embodied if some properties of E are processed in the same way as the properties of one's body". In de Vignemont's definition, one may feel a tool – such as a hammer – to be embodied without feeling that the tool is part of one's body (no ownership). Kilteni and Slater [Kilteni et al., 2012a] adapt de Vignemont's definition considering the Sense of Embodiment (SoE) in the context of VR: "SoE toward a body B is the sense that emerges when B's properties are processed as if they were the properties of one's own biological body". Here we make a distinction to the use of the expression "Sense of" before Embodiment. As described by de Vignemont, "sense of" refers to the fact that the subject feels such phenomena, instead of only knowing it exists/happens. For instance, one may learn and believe due to anatomy studies that she has a gallbladder that is part of her body, but we do not feel the gallbladder as being ours. However, the way we experience our relation with limbs such as arms and hands are more complex and complete than that, as we not only know we have arms and hands, but we feel and control them. In other words, it is the subjective experience of embodiment – of higher interest from the VR perspective – in which a healthy subject may be deceived to accept and believe in a virtual representation attributed to her.

According to [Longo et al., 2008, Kilteni et al., 2012a], a successful sensorial manipulation of the sense of embodiment (SoE) may rise from three components of the instantaneous bodily self:

- Sense of agency: [Sanchez-Vives et al., 2010] defines it as the sense of "global motor control, including the subjective experience of action, control, intention, motor selection and the conscious experience of will". It is proposed that the sense of agency emmerges from the comparison between predicted and actual sensory consequences of one actions [David et al., 2008].
- Sense of Body ownership: it refers to one's self-attribution of a body [Blanke, 2012, Gallagher, 2000]. The emmergence of a sense of ownership is said to rely on the spatial and temporal correlations among sensory cues (visual, tactile, proprioceptive, vestibular) that arise from our body [Jeannerod, 2004].
- Sense of self-location: is determined by a certain volume in space, where one feels to be located [Kilteni et al., 2012a]. Under normal circumstances, one experience to be located within one's own body. However, this unity may be disrupted under certain circunstances, such as when a person have an out of the body experience (OBE) [Lenggenhager et al., 2006].

Thus, SoE refers "to the ensemble of sensations that arise in conjunction with being inside, having, and controlling a body" [Kilteni et al., 2012a]. It is also argued that embodiment is not a discrete condition, that is, it can be experienced in different levels, depending on how many properties are met as well as the intensity that they are met.

2.2.3 Manipulating the Bodily Self

The instantaneous experience of embodiment emerges from the coherent multisensorial integration taking place in the brain, which has been referred to as bodily self-consciousness (the pre-reflective sensation of being the subject of an experience) [Legrand, 2006, Blanke, 2012, Blanke et al., 2015]. Experimental protocols have shown that this body representation is much more malleable than commonly assumed. For instance, conflicting sensorial stimulation can temporarily change how one perceives properties of their own body (i.e. an altered bodily self-consciousness). Notably, it can lead to the illusion of owning a fake – either material or virtual – limb [Botvinick and Cohen, 1998], body [Slater et al., 2010b], and even another individuals' body [Petkova and Ehrsson, 2008]. These illusions are explored below.

Rubber Hand Illusion

In the rubber hand illusion (RHI) a fake model of the hand is placed at a position coherent with the subject body, while a physical barrier is used to occlude the real hand from sight. The experimenter then repeatedly and synchronously strokes both, the real and the rubber hand. By watching the fake hand being touched at the same moment and region as felt by the real hand, the subject may experience and report the sensation that the fake hand belongs to her body, and even despond to threats directed to the fake hand [Armel and Ramachandran, 2003]. Additionally, when asked to use the opposite hand to point to where the hidden hand is, subjects tend to wrongly localize the position of the occluded hand towards the rubber hand. This measurement is better known as *proprioceptive drift*, and correlates with the subjective report of ownership provided by the subjects. Figure 2.2 illustrates the RHI.

The protocol above describes a bodily illusion induced by the congruent visuo-tactile stimulation, but a variaty of experiments explored other senorial congruences, such as sensorimotor correlations. For instance, passive and active synchronous movements (visuo-proprioceptive and visuo-motor) were shown to elicit the illusion [Tsakiris et al., 2010, Walsh et al., 2011, Kalckert and Ehrsson, 2012a, Kalckert and Ehrsson, 2012b], possibly stronger for active movement [Tsakiris et al., 2010]. Moreover, interoceptive signals are also known to play a relevant role, modulating the intensity [Tsakiris et al., 2011] and even driving the illusion [Suzuki et al., 2013].

Nonetheless, the RHI was shown to work in VR setups. It has been induced through visuo-tactile [Slater et al., 2008] and visuo-motor [Sanchez-Vives et al., 2010, Yuan and Steed, 2010] synchrony.

The RHI is generally not successful if the sensorial stimuli applied to real and fake hands are asynchrounous, if the fake hand is placed at a position incongruent or too far relative to the body, or if the object replacing the hand does not resembles a hand [Tsakiris and Haggard, 2005, Blanke et al., 2015].

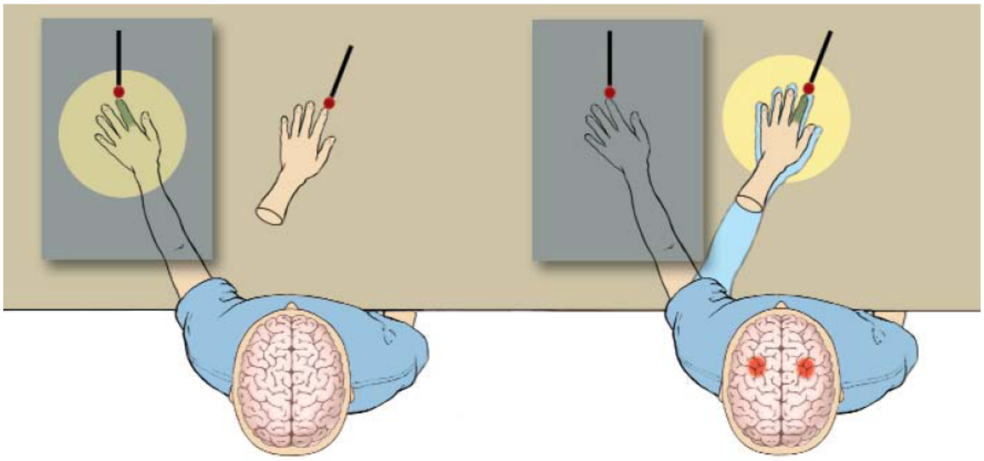


Figure 2.2 – Illustration of the rubber hand illusion. After synchronous and repeated stroking of the real hand and the rubber hand, the subjects feel as if the seen fake is theirs own. Evidence shows that visual receptive fields are brought into allignment with the rubber hand (yelow areas), and that visual information realigns the proprioceptive map (blue arm profile). Figure from [Botvinick, 2004].

Full Body Ownership Illusion

The full body ownership illusion (FBOI) is analogous to the RHI. It has been studied using cameras and VR, demonstrating that a whole alien body (real, fake or virtual) can be felt as ones' own body. In [Petkova and Ehrsson, 2008], Petkova and Ehrsson demonstrate two setups capable of inducing the sense of ownership of a mannequin and even of another person's body. In the mannequin protocol, a camera is positioned on the head of the surrogate body, and the image of the camera is transmitted to an HMD worn by the subject. By applying synchronous visuo-tactile stimulation to the abdomen of the subject and the mannequin body (Figure 2.3), subjects reported the feeling of ownership toward the mannequin. Additionally, when the experimenter slides a knife over the abdomen of the mannequin, a strong physiological response was observed. This response was stronger when the visuo-tactile stimulation was applied synchronously. Figure 2.3 illustrates this illusion.

Similarly, [Slater et al., 2010b] have replicated these results in VR. With visuo-tactile stimulation and visuo-motor congruence of the point of view, the experiment led male subjects to feel ownership of a young female body.

2.2.4 Perpective taking

Alterations to the sense of embodiment can be observed even when the surrogate body position does not coincide with the point of view of the scene, i.e. seen from a third person perspective (3PP). In this setup a stereoscopic camera watches a subject or mannequin from the back, and transmit the 3PP image to the subject through a HMD. The synchronous visuo-

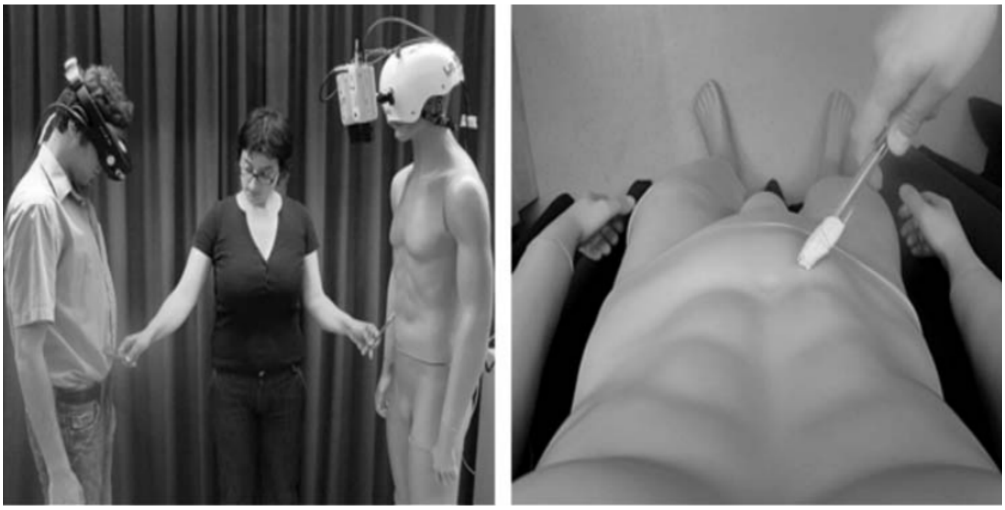


Figure 2.3 – Full body ownership illusion. After synchronous and repeated stroking of the real body and the dummy body, subjects feels as if the seen fake body is their own body. There is an associated increase in skin conductance response when the dummy body is threatened by a knife (Figure from [Petkova and Ehrsson, 2008]).

tactile stimulation delivered to the subject or mannequin back (Figure 2.4) was shown to provoke the sensation of ownership of the distant body, seen from a third person perspective [Lenggenhager et al., 2007]. These results have been replicated in [Lenggenhager et al., 2009].

However, literature diverges with respect to FBOI in 3PP. Other experiments suggest that 1PP plays a major role. For instance, [Slater et al., 2010b] performs an experiment including visuo-tactile congruence and perspective as factors. Their results suggest that 1PP is a critical factor for the ownership of a body transfer illusion, contrasting previous studies that suggest visuo-tactile congruence to be the main contributory factor to the ownership illusion. Moreover, [Pomés and Slater, 2013] presents a setup in which a virtual body is seen from behind. Congruent or incongruent visuo-tactile stimulation could be applied, and the subject could control the arms of the virtual body. No effect of visuo-tactile congruency has been found, and the ownership scores were generally low.

While we agree that perspective might significantly impact the sense of ownership, in Chapter 4 we sought to further this knowledge by manipulating perspective and visuo-motor congruence of the whole body in two experiments. We also highlight that the animation algorithms used in [Pomés and Slater, 2013] are simpler and produce more artifacts than the ones we use here.

From a more practical standpoint, changing the perspective from first (1PP) to third person perspective (3PP) allows taking a new and potentially more informative point of view within a VR application (such as for training [Maupu et al., 2009, Salamin et al., 2010, Covaci et al., 2014]).

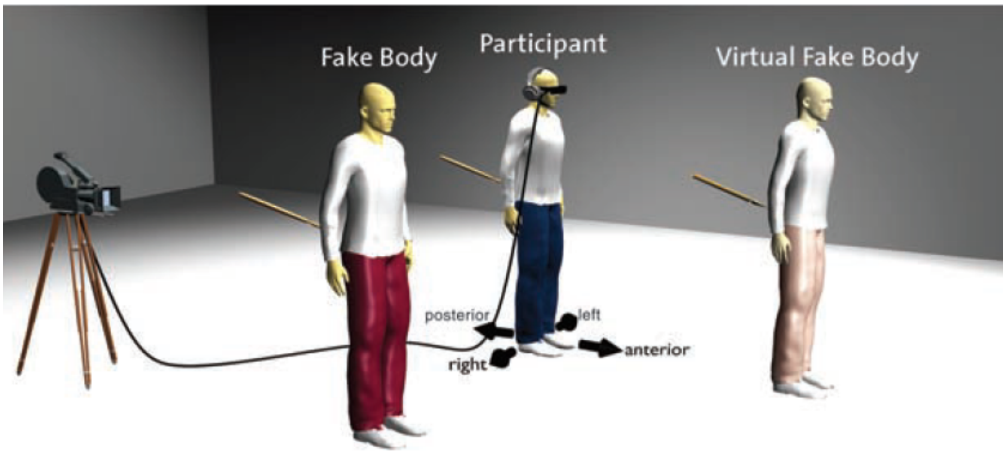


Figure 2.4 – Full body ownership illusion in 3PP. After synchronous and repeated stroking of the real body and the dummy body, subjects feels as if the seen fake body is their own body. This illusion is argued to be valid even if the body is seen from a third person perspective (Figure from [Lenggenhager et al., 2007]).

2.3 The self as an actor

The movement distortion we propose in Chapter 5 is designed to guide reaching movements, facilitating or hindering the completion of a reaching task. Thus, we are especially interested on the body of knowledge that describes how motor planning and movement control are managed.

Moreover, the ability to differentiate self-generated actions from externally generated stimuli is transparent to the healthy subject. Although we take this ability for granted, empirical evidence shows that the mechanisms that are likely to control the self-attribution of actions have to undergo constant adaptation. Here we briefly discuss these mechanisms.

2.3.1 Internal Models of movement control

An internal model is a system that mimics the behavior of a natural process, predicting the future state of a system, such as the velocity and position of an acting limb [Wolpert et al., 1995]. It is widely accepted to explain the representation of movements and intentions. The execution of an internal model does not imply an external movement, it can also result from imagining an action or seeing an actions, such as in Ramachandrans mirror-box [Ramachandran and Rogers-Ramachandran, 1996].

It is argued that the internal model controlling the movement of a limb makes use of an inverse and a forward model of the limb [Desmurget et al., 1999, Desmurget and Grafton, 2000] (Figure 2.5). The inverse model translates an intention into motor commands (e.g. intended end effector position into rotations for the chain of joints), while the forward model of the limb uses a copy of the motor commands (known as the efference copy) to predict the sensory consequences for that motor command [Wolpert et al., 1995]. The predicted consequences are further compared to the actually sensed information of that movement (reafference), a self generated movement is expected to result in minimal discrepancy, and its sensorial consequences are suppressed [Kawato, 1999, Wolpert and Flanagan, 2001]. Notably, Helmholtz discusses the need for such mechanism to explain why the image of the world is perceived still following the voluntary movement of the eyeball [von Helmholtz, 1910]. In comparison, if one gently taps the eyeball with the finger, the whole world seems to move. Therefore, the intention to move and the motor commands provided to the extraocular muscles might be responsible for this difference in perception. This mechanism is also used to explain why people usually cannot tikle themselves. In [Blakemore et al., 1999], the increased latency between one's action and its sensory consequences (mediated by a mechanical device) resulted in a higher sensation of "tickliness", despite the fact that subjects were not aware about the manipulation of latency.

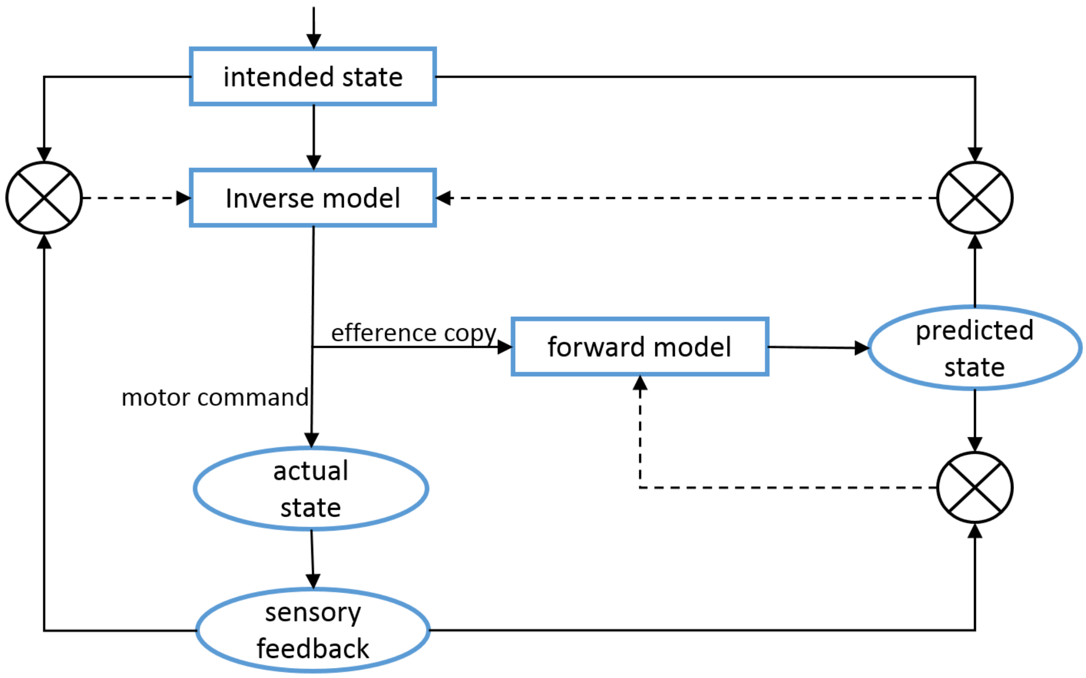


Figure 2.5 – Internal model of action. The inverse model is used to plan an action, and the forward model is used to monitor this action. The output of comparisons (dotted lines) are used to refine the inverse and forward models (motor adaptation). According to [Frith et al., 2000, Blakemore et al., 1999, Blakemore et al., 2002], these comparisons also play a role in the attribution of action (agency). Figure adapted from [De Vignemont and Fourneret, 2004].

2.3.2 Comparator model and the sense of agency

The comparison mechanisms described by internal models have been extensively suggested to underlie the self-attribution of actions [Blakemore et al., 2002, De Vignemont and Fourneret, 2004, Frith, 2012]. In short, this view suggests that if significantly incongruent signals arrive to one of the comparators (Figure 2.5), the subject may become aware that something went wrong with an intended action. [Blakemore et al., 2002] states: "we seem to be unaware of the results of the comparison between the predicted and intended outcome of motor commands, and the comparison between the predicted and actual sensory feedback, as long as the desired state is successfully achieved". Under this view, the comparator mechanism has a central role in segmenting self from externally generated actions. This proposition could model well certain clinical conditions known to impair the sense of agency, such as delusions of control in schizophrenic patients [Franck et al., 2001, Blakemore et al., 2002].

Moreover, in a pioneering study [Nielsen, 1963] has demonstrated that, when the visual feedback of the movement of a healthy subject is replaced by the movement of a second person, the subject might miss-attribute the seen movement to himself. This was the case even when there was a discrepancy between the performed and seen movements, with subjects reporting the feeling of strangeness or impression that their hands have been pulled by some external force. This experiment illustrates that there is a limit within which a person is unaware about discrepancies that may occur to her own movements.

2.3.3 Sensorimotor adaptation

Sensorimotor adaptation is essentially an iterative process of optimization of the inverse and forward models by minimizing discrepancies between predicted and actual outcome of actions. Motor adaptation is essential to successfully interact with the external world and to accommodate to the constant body reshaping that people undergo. For instance, forward models are only useful if they can produce unbiased predictions, thus a forward model has to remain calibrated through motor adaptation, and when the discrepancy between estimated and actual sensorial feedback configures an error, this error can be used to improve the forward model [Shadmehr et al., 2010].

Notably, [von Helmholtz, 1910] describes the adaptation effect following the use of prism goggles, which rotate the whole field of view of a subject by a fixed angle. He demonstrated that a subject can adapt her movements rather quickly to comply with the altered visual feedback of space. At the start of the exposure, the subject commits errors when pointing at targets. After an adaptation period, the subject becomes capable of compensating for the angular displacement. Following the exposure period, an after effect is observed, in which the subject tends to commit pointing errors to the opposite direction, quickly restoring to the correct visuo-motor mapping.

Moreover, motor adaptations may happen even if the subject is unaware about the manipulation. For instance, [Kannape and Blanke, 2012, Kannape and Blanke, 2013] shows that when subjects are faced with an angular or temporal deviation of their gait, they tend to adjust for this deviation without being aware of such manipulation (up to a certain threshold).

In the experiment we present in Chapter 5, we quantify by how much we can distort the visual feedback of a reaching movement. Relating our experiment to the concepts we discussed here, we want to know to what extent subjects will correct their movement in order to comply with the distortion, before becoming aware that the seen movement does not correspond with their actual movement.

3 Increasing the Awareness of the Virtual Environment

In this chapter we discuss two approaches designed to provide increased awareness of a controlled virtual body posture and its relation with the environment. In our scope, to be aware means to be informed through perception about the objects, conditions and events involving the user in the virtual environment. In the first approach we use non-planar projections as a means to increase the Field of View in embodied Virtual Reality (Section 3.1). In practice this requires renouncing perspective projection in favor of a non-planar one. In the second approach we examine how first and third person perspectives could be combined in order to sum up their advantages (Section 3.2). Namely, 1PP of a virtual body can consistently induce the sense of ownership of the surrogate body, while 3PP can provide constant feedback of the virtual body posture and its relation with the environment.

By exploring these approaches we expect to better understand how the sight of the virtual body affects the sense of embodiment, virtual body/environment relation and quality of interaction. We close the chapter with an overview of the techniques, and by pointing the one of our preference for a more complete evaluation (presented in Chapter 4).

3.1 Non-planar Projection with HMDs

This section has been adapted from [Debarba et al., 2015b].

Even with the recent rise of affordable HMD, which brought VR back into the popular imaginary while approaching the mass market, delivered FoV is still a lot inferior than the human eye FoV. Oculus DK2, which is arguably among the most popular models offers a maximum of $\approx 106^\circ$ FoV, even though in practice most users experience something around 90° as the FoV also depends on the eye/lenses and lenses/screen distances. On the other hand, human FoV is $\approx 180^\circ$ horizontally, and $\approx 135^\circ$ vertically (with a downward bias). This allows us to be visually aware of our body at all times –the fact that we have one, its posture, and its relation with the environment–, which is not the case while using an HMD and controlling a virtual body.

We explore the use of non-planar projections to address this limitation, i.e. showing more of

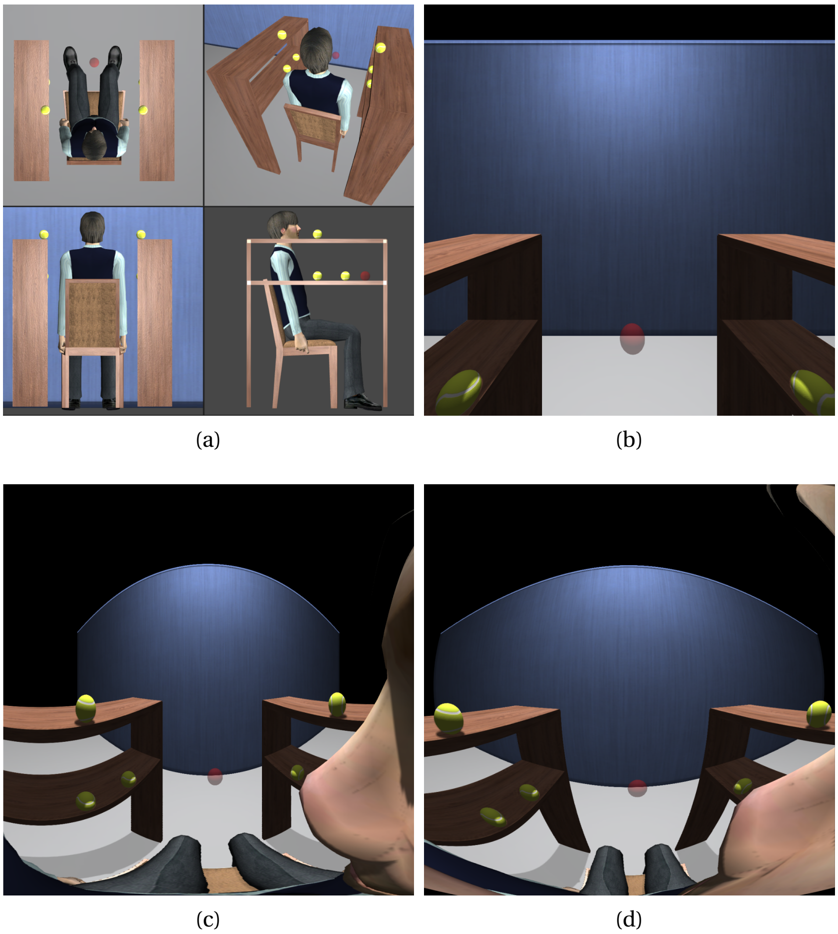


Figure 3.1 – Virtual environment and screenshots of tested projections. Orthogonal and perspective views of the virtual environment used in the experiment (a). Render capture of the left eye for Perspective (b), Equirectangular (c) and Hammer (d) projections respectively. Notice that it is possible to see the virtual body (including the nose) with the non-planar projections.

the virtual body and environment at the cost of altering the projection we experience in the natural world. To assess this issue, we performed an experiment that evaluates how one feels and performs in a selection and docking task while controlling a virtual body in a cluttered virtual environment. Our implementation and choices of non-planar projections notably relies on cartography, which studies means to represent the surface of the Earth over a plane. Here we compare the regular Perspective projection ($\approx 106^\circ$ vertical FoV), with the Hammer ($\approx 180^\circ$ vert. FoV) and Equirectangular ($\approx 180^\circ$ vert. FoV) non-planar projections (Figure 3.1bcd). During the development phase, we tested several non-planar projections and selected Hammer and Equirectangular, being consistent with the positive results presented in [Ardouin et al., 2013].

Equirectangular is a cylindrical projection with the property of equidistance, while Hammer is an azimuthal projection with the property of equal-area. Detailed information on these projections can be found in [Kennedy and Koop, 1994].

In order to assess whether seeing the virtual body more often increases subject's identification with it, we evaluate the reported sense of Embodiment with a questionnaire considering: (a) sense of agency, i.e. feeling of control over the virtual body; and (b) sense of body ownership, i.e. feeling that a virtual body is one's own body. The questionnaire was based on the one presented in [Slater et al., 2010b]. A simulation sickness questionnaire (SSQ) was also administered [Kennedy et al., 1993]. The questionnaires we used are available in Appendix C

3.1.1 Related Work

Non-planar projection has been studied as an alternative to perspective projection in order to increase the FoV of a camera (i.e. perspective is limited to $< 180°$), and/or to keep better proportionality for the information:rendering area ratio (i.e. large FoV perspective projection tends to render most of the information in a small region at the center of the image). This allows for the presentation of more environmental information with reduced effort (e.g. less camera movements). Early works have approached non-planar projections on conventional displays, while more recent work have also explored this matter in an HMD and augmented reality context.

[Glaeser and Gröller, 1999] proposed the use of such mappings as an alternative to perspective rendering when a wide FoV is required. When rendering in a desktop, they argue that the distortions introduced to the image by a non-planar projection are less detrimental to its comprehension than the distortion due to a wide FoV perspective projection. On a related investigation, [Ardouin et al., 2013] evaluates how subjects perform in a navigation task for various 360° non-planar projections, with the VE seen through a 22 inches screen. They point to some advantages and favored the usage of Hammer and Equirectangular projections. Furthermore, [Mulloni et al., 2012] explores how different panoramic images in a computer screen affect subjects' ability to find and correctly point to objects at their surroundings.

Closer to our context, a few studies also approached non-planar projections rendered by HMDs. [Ardouin et al., 2012] proposed a system that delivers 360° of horizontal FoV to the user. The image is captured by a camera from the top of the users head and is fed to the HMD, providing easy and intuitive control of the point of view. However, the project is mostly conceptual and the evaluation was solely based on user impressions. [Orlosky et al., 2014] brings a deeper study using a pair of cameras and 233° fisheye lenses to evaluate perception of objects in the periphery of vision.

To the best of our knowledge, no past work explored the use of non-planar projections in an immersive VR setup and from the perspective of embodied interaction.

3.1.2 Materials and Methods

Projection Implementation

The experiment was developed with the Unity game engine. To obtain the non-planar projection images, we render a 360° image using six 90° perspective cameras. We then map the 6 rendered images into a high density cube mesh. This cube mesh is then transformed into a sphere by normalizing the length of vectors formed by each vertex relative to the center of mass of the whole mesh. Finally, the resulting spherical mesh is transformed into a plane with a cartographic projection equation (latitude and longitude coordinates into x and y coordinates). This approach is similar to the one employed by [Bourke, 2009] to project over a hemispherical surface.

Virtual Environment and Task

The VE consisted of a chair, on which the virtual body is seated. Additionally, a pair of two levels shelves were placed along each side of the chair, where targets could appear at 6 predefined positions (2 targets within and 4 beyond a 90° FoV while looking forward). The target has the shape of a tennis ball with 6.7 cm diameter. An additional docking volume is shown in front of the virtual body, it has the same size as the targets and is rendered with transparency. An overviews of the VE is shown in Figure 3.1a.

The task consisted of reaching and docking targets with the dominant hand (as indicated by the subject), in each trial: a target appears; the subject searches for it; then (s)he moves the end effector in order to intersect the target; (s)he selects the target by holding a trigger button and translates it toward the semi-transparent docking volume; finally, (s)he releases the target by releasing the trigger button; the next target appears. Subjects were asked to perform the task as fast and accurately as they could. There were a total of 6 different targets, which were repeated 4 times each, for a total of 24 *trials* per *block*. Subjects performed a *block* for each tested projection, yielding a total of 3 *blocks*. If no interpenetration between target and docking volume occurs (i.e. docking error bigger than 6.7 cm), the trial is marked as failed and has to be repeated by the end of the block.

Tracking and Motion

A *PhaseSpace ImpulseX2* with 14 cameras is used to track the position of 4 LED markers attached to hands and elbows. To retrieve hands orientation and allow for input selection, a pair of *PS Move* controllers are used. They communicate with a *Playstation 3*, which uses the software *Move.me* to stream the controllers data to our program. The arms of the virtual body are driven by inverse kinematics, which defines a posture relying on the 6 degrees of freedom of the tracked hands as well as the position of the tracked elbows. LED markers were also added to the chair for a calibration step (match real and virtual chairs). The subjects were asked (and assumed) to keep their posture and avoid moving trunk and legs – a predefined seated posture is applied to the virtual body for these body parts. For the experiment, a gender matching virtual body was used, and its height was scaled to match the height of the subject.

An *Oculus DK 2* is used as display, the Oculus also provides drift-free orientation based on its inertial sensors and optical tracking. This orientation is used to rotate the virtual camera and the head of the virtual body. No position tracking was used, thus the point of view rotated around a predefined pivot position in the virtual body neck. The setup is shown in Figure 3.2.

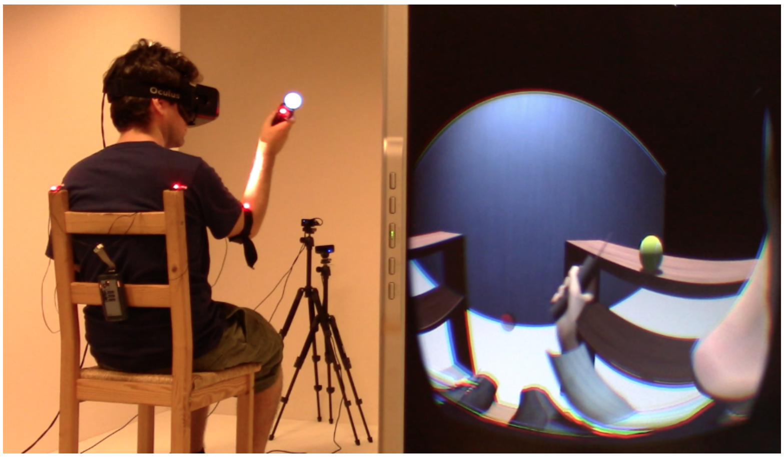


Figure 3.2 – Overview of the system setup. An Oculus DK2 HMD was used as display, and to track head orientation. Optical markers were used for position tracking, while a pair of PSmove controllers were used to track hand orientation and to acquire the target.

Experiment Design

The experiment followed a within-subject design with random projection order. A total of 6 subjects aged from 17 to 25 participated on the experiment (1 female). *Projection* was the main factor to be controlled. *Target position* was also treated as a factor for the time related responses.

3.1.3 Results

Time and precision: we consider reach time (RT) – time until the selection of a target –, dock time (DT) – time from selection to dock –, dock error (DE) – error in cm between predefined docking position and actually docked position. These values are computed by taking the median of the successful trials for each combination of *subject*, *target position* (except for DE) and *projection*. Subject 2 had difficulty to adapt to the non-planar projections and is not considered for RT, DT and DE analyses. Subject 2 had to perform 104 trials to successfully dock a total of 48 targets (i.e. docking error of less than 6.7 cm) for Equirectangular and Hammer projections together. Other subjects had to repeat a maximum of 4 trials during the whole experiment.

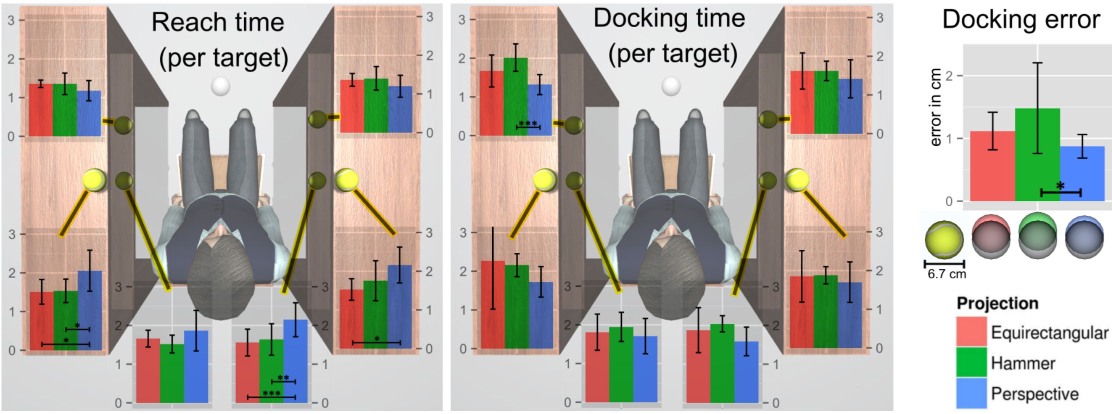


Figure 3.3 – Non-planar projections performance results. Time to reach (left) and dock (center) in seconds, and docking precision (right) in centimeters. Reach time with non-planar projections was significantly shorter for most of the lateral targets. Perspective presented time advantage for docking , as well as smaller docking error. Error bars indicates the estimated standard deviation. Significance results were computed with Tukey HSD test. '$*$', '$**$' and '$***$' indicates $p < .05$, $p < .01$ and $p < .001$ respectively.

For RT and DT, statistical analysis was carried with two-way ANOVA with *projection* and *target position* as independent variables. RT presented an interaction between *projection* an *target position* ($F(10,40) = 3.9, p < 0.001$), indicating a trade-off for these factors. For DT only target position yield a significant difference ($F(5,20) = 6.848, p < 0.001$). We further analyzed the effect of *projection* to RT and DT per *target position* using Tukey HSD test. For DT, difference was significant for one of the frontal targets in favor of Perspective as compared to Hammer. For RT, the difference was significant for 3 of the lateral targets in favor of Equirectangular as compared to Perspective, and for 2 of the latter targets in favor of Hammer as compared to Perspective (details presented in Figure 3.3). Further analysis with Tukey HSD to the mean head movement per trial (in radians) shows that subjects performed significantly less head turns with Equirectangular ($M = .57$ $SD = .11$) and Hammer ($M = .69$ $SD = .16$) projections as compared to Perspective ($M = .83$ $SD = .10$, with $p < .001$ and $p < .011$ respectively). Suggesting that subjects took advantage of the increased FoV, which in turn led to a decrease in search time.

To analyze DE we used Tukey HSD corrected for multiple comparisons with *projection* as the only independent variable, which has shown that Perspective performed significantly better than Hammer. Figure 3.3 shows the RT, DT and DE means and standard deviations for each projection.

Collisions with environment: we assess subjects' understanding of the environment and its relation with the virtual body by considering the mean of collisions per trial (MCol) between virtual body and shelves while performing the task. We compute MCol for the arms (upper arm and forearm), hands, and arms/hands together. If simultaneous collisions happens, only one collision is considered. Analysis was carried with Wilcoxon signed rank test with

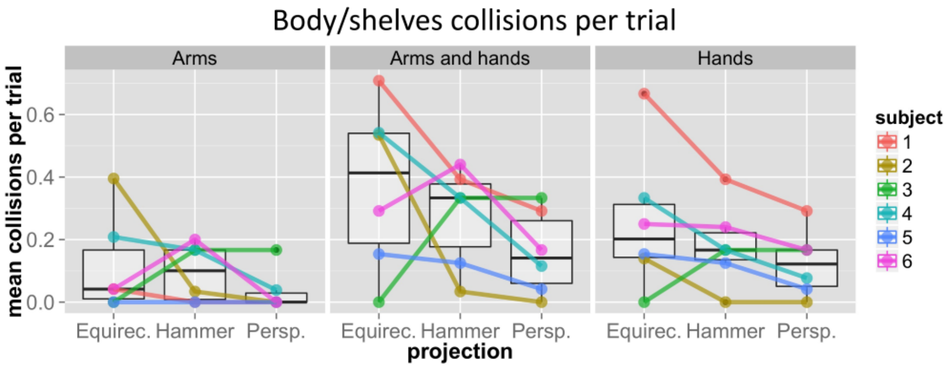


Figure 3.4 – Collisions of the body with shelves per subject. Everyone but subject 3 presented less body/shelves collisions while using perspective projection.

Holm-Bonferroni correction. The increase in collisions with Hammer and Equirectangular projections were not statistically significant when compared to Perspective projection (p = 0.059 for corrected α = .0167 and p = .156 for corrected α = .025 respectively). MCol per subject for each *projection* is shown in Figure 3.4.

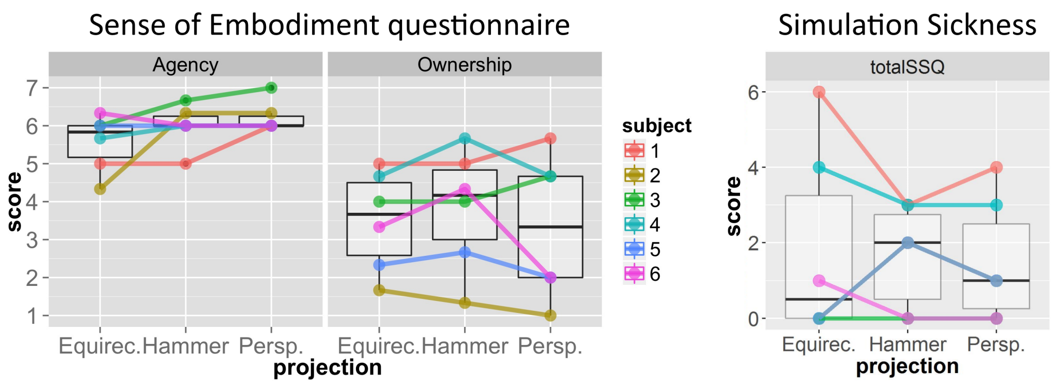


Figure 3.5 – Reported sense of agency and sense of body ownership, and simulation sickness questionnaire results.

Sense of embodiment: we assess the senses of Agency (AG) and Body Ownership (BO) with a questionnaire. Agency score is the mean of 3 questions, whether the subject felt: to be in control of the VB, not to be in control of the VB, and the VB was responsive to his/her movements. Ownership score is the mean of 3 questions, whether the subject felt: that the VB was his/her body; that the VB was not his/her body; to be wearing VB clothes. Negative questions had their score inverted before taking the mean score.

Statistical analysis was carried with Wilcoxon signed-rank test and Holm-Bonferroni correction. No significant difference was found for both agency and ownership (smaller p-value at p = .134 for corrected α = .0167). As the virtual body could be seen the whole time, we were expecting an increase with non-planar projections. However, these results suggest that the non-planar projections had no consistent effect to the reported sense of embodiment (Figure

3.5).

Cybersickness: was assessed with the simulation sickness questionnaire (SSQ). The mean score for Equirectangular, Hammer and Perspective projections were 1.83, 1.67 and 1.5 respectively (Fig 3.5). No statistically significant difference was found with Wilcoxon signed-rank test and Holm-Bonferroni correction (smaller p-value at $p = .572$ for corrected $\alpha = .0167$).

3.1.4 Discussion

On the one hand, our results indicate that the large FoV non-planar projections may increase reaching performance when interacting with targets beyond the perspective FoV, leading to less head rotations. On the other hand, the time needed to dock as well as the docking precision were reduced, even though the difference was not always statistically significant.

Additionally, the amount of collisions between the virtual body and environment obstacles have increased with the non-planar projections. Although the difference was not significant, we expected that the visual feedback of the body would improve subject's perception of the virtual body/environment relation, which was not the case. Even with the reduced FoV, the spatial model [Tversky, 1993] that one creates in perspective projection seems to be more effective than the more constant – but distorted – visual feedback that the large FoV non-planar projections provide. This did not prevent subjects from succeeding in the proposed experiment, but may point a poor translation of structural visual information into movement planning, such as observed by the usage of a prism glasses [von Helmholtz, 1910]. Alternatively, the distorted feedback may negatively affect the recalibration of proprioception by integrating sensorial input into inaccurate postural information.

Finally, the non-planar projections were not detrimental to the reported senses of agency and of body ownership. They also did not elicited significantly stronger cybersickness as compared to Perspective projection.

We were initially expecting to find differences between Hammer and Equirectangular projections. More specifically, we had the expectation that Equirectangular projection would provide more accurate structural information, given that it preserves some straight lines (Figure 3.1c). Such information could be useful to prevent collisions and increase docking precision. As both non-planar projections presented very similar performance, this hypothesis could not be verified. As a matter of fact, our results are in line with [Ardouin et al., 2013], in which navigation performance and subjective evaluation ranked Equirectangular and Hammer projections to be very close as compared to the other projections the authors have tested.

Furthermore, there are other basic 3D user interaction tasks that should be considered in the future, such as navigation and finer manipulation. It might be the case that non-planar projections are less suitable for fine manipulation of objects (e.g. involving the fingers). In addition, SSQ scores may be altered in a navigation task, given that change in visual flow promoted by non-planar projections would be coupled with additional forms of movement.

Finally, we point to the fact that subjects only used each projection for 2 to 5 minutes in our experiment. A long term adaptation might play a strong role on performance, and no related work explored this venue yet.

3.2 Combining 1PP and 3PP

Part of this section has been adapted from [Debarba et al., 2013].

In this section we build an argument in favor of combining 1PP and 3PP. We emphasize how this could bring together the desirable characteristics we exposed in the introduction chapter, namely, strong sense of embodiment and awareness of the virtual body posture and its relation with the virtual environment. We propose and discuss a design space of perspective combinations. The evaluation of our favored perspective combination interface is later presented in section 4.3.

Carefully designed third person perspective (3PP) is often used in games to convey information about user surroundings [Maupu et al., 2009, Taylor, 2002], information that otherwise would not be observed in first person perspective (1PP) due to the narrow field of view (FOV) of regular screens. On the contrary, the human horizontal FOV is close to 180°, hence allowing us to easily spot events involving our body. Immersive displays such as the CAVE or head mounted displays (HMD) may reduce the pointed visual issue for 1PP, but VR is still limited when rendering information to other senses involved on full-body interaction. For example, haptic displays may be mature for interaction with probes and even hands, but are still unable to provide realistic full-body haptic rendering.

Nevertheless, full-body interaction became recently accessible with the advent of Kinect. Notably, full-body interaction games tended to adopt the 3PP instead of 1PP view of the controlled avatar, demonstrating user's need for continuous feedback of the character pose and of its relation with the environment. Many factors yield the use of 3PP visualization for full-body interaction in ordinary video game setups: the narrow FOV of a common screen; the latency of the tracking system; the mismatch in scale, position, orientation and articulation model between real and virtual body, limbs and joints. On the other hand, all these limitations do not prevent users from performing the desired pose and successfully interacting with the environment when using the 3PP. This exalts 3PP potential to convey body posture awareness.

Therefore, we propose to combine 1PP and 3PP, taking advantage of their qualities in order to complement each other. With 1PP users can experience stronger sense of embodiment towards the avatar [Blanke, 2012, Havranek et al., 2012, Petkova et al., 2011, Slater et al., 2010b]. On the other hand, with the 3PP users can be constantly aware of the virtual body pose [Maupu et al., 2009] and its relation with surrounding VE [Boulic et al., 2010, Taylor, 2002].

With respect to our needs, we argue that the best combination is to use the 1PP as the main viewpoint, and to augment it with 3PP. This approach supports the aforementioned advantages of each perspective, which are deeply linked to the general goal of this thesis.

3.2.1 Related Work

There are several studies employing multiple viewports for to visualize/navigate/interact in/with VE [Stoakley et al., 1995, Schmalstieg and Schaufler, 1999, Kiyokawa and Takemura, 2005, Hirose et al., 2006, Wang et al., 2011]. However, only one addresses the issue of embodied interaction and body relation with the environment [Salamin et al., 2008]. Additionally, a perspective comparison paper also speculate on this possibility [Maupu et al., 2009].

A classical approach is the worlds in miniatures [Stoakley et al., 1995], which augments the 1PP visualization of a VE using a dynamic viewpoint for the overview and manipulation of the VE. [Schmalstieg and Schaufler, 1999] proposed SEAMS, which allows the user to observe and move to different scenes, it conceptually seams worlds together through viewports. [Kiyokawa and Takemura, 2005, Hirose et al., 2006] proposed a multi-viewport system for navigation and manipulation of remote objects, simultaneous viewports are used to visualize and to give interactive access to content that may be in another VE or out of users visual reach. [Wang et al., 2011] proposed pop-up depth views, in which perception of a 3D cursor is enhanced by the use of orthogonal viewports. These viewports are positioned around the cursor, and it is proposed that they pop-up whenever cursor movement speed is below a certain threshold.

Perhaps, the setup proposed by [Salamin et al., 2008] is the closest to our proposition. It equips a subject with a HMD and two cameras, one fixated to the HMD (1PP) and a second to a metallic structure attached to a backpack (3PP). The image of the 3PP camera is fed to the HMD, while the image of the 1PP camera is used to overlay the body of the subject in the 3PP image. It is proposed that this could reduce the impact of body occlusion present in 3PP.

Additionally, [Maupu et al., 2009, Boulic et al., 2009] compare 1PP with two orthogonal 3PP (from the left and from behind the region of interest). It assessed the performance of these two conditions for a reaching task, suggesting the adoption of a 3PP to control the posture of a virtual avatar and its relation with the environment. Indeed, they conclude by proposing a combined interface that could alternate across perspectives, in which the user would be able to switch to 3PP to adjust his avatar body posture, and return to 1PP to experience its point of view at that given pose.

This work focuses in a yet overlooked problem. How to combine 1PP and 3PP while preserving the sense of embodiment of the virtual body. We propose a design space to classify these interfaces and to guide their design.

3.2.2 Design Space for Combining Perspectives

There is not an obvious single manner to combine 1PP and 3PP, in fact, the possible combinations are endless, and a single setup is unlikely to generalize well across different applications.

To address this matter, we formulate a design space for how the combination of 1PP and 3PP could be carried (Figure 3.6). Based on this design space, we argue that one has to answer the following design questions to define a combination proposition:

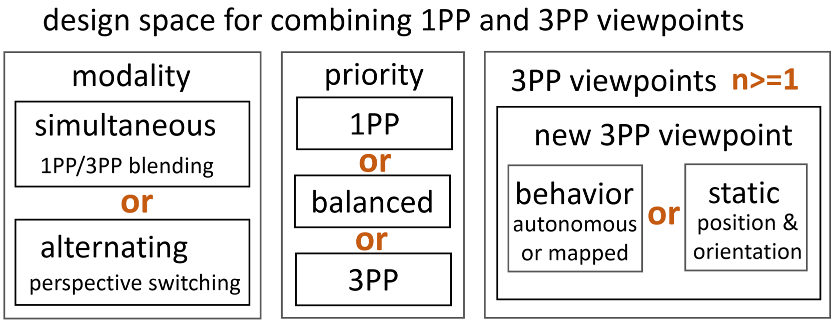


Figure 3.6 – Design space for combining 1PP and 3PP viewpoints. To design an interface one has to decide: (i) if 1PP and 3PP will be presented simultaneously or in alternation; (ii) if one of the perspectives will be enforced by the interface; (iii) the behavior and properties of 3PP cameras; (iv) the number of 3PP cameras.

(i) Will 1PP and 3PP be presented simultaneously or in alternation?

In the simultaneous presentation of perspectives, we additionally have to decide how to blend the perspectives. This could be achieved either by overlaying information of the scene or by placing extra viewport(s). For instance, in Figure 3.7 we explore the placement of additional viewports, associating them to real life metaphors. In the specific setup depicted, 1PP is the main viewpoint through which a subject experience the VE, and 3PP is inserted in the scene through viewports (See Figure 3.8). We detail this setup and its implementation in Section 3.2.3 below.

On the other hand, perspectives may be presented in an alternating manner. In this case, one has to consider the most effective way to switch between perspectives, so that it is time efficient and it does not disturb the VR experience (e.g. preventing interpenetration and cybersickness). We make considerations on the design of an alternating interface in Section 3.2.4.

(ii) Will one of the perspective options be enforced over the other by the interface (e.g. be the default mode, appear more often, etc)?

In our specific case we put high priority to the sense of embodiment, thus we tend to favor 1PP as the main point of view. This is the case for the simultaneous perspective combination proposed in section 3.2.3. However, the alternating interface we propose does not enforce an specific perspective, as it allows the subject to switch to what she believes to be the most appropriate at any given moment (Section 3.2.4). Alternatively, [Salamin et al., 2008] presents a setup in which 3PP exerts the main role, and the 1PP image is presented overlaying part of the user body, conveniently addressing the issue of body occlusion in 3PP.

(iii) Should the 3PP cameras have an autonomous behavior, a static placement, or be directly controlled?

While the settings of 1PP are straightforward, with a unique assumed placement (the virtual body eyes) and a behavior expected to mimic subjects' head movements, 3PP has no well defined constraints. 3PP could be programmed with an autonomous behavior, such as following the virtual body and presenting information deemed critical by the controlling algorithm, it could also be directly mapped to be under subjects' control, or it could be strategically positioned to support a given task.

Based in our observations and assumed real life experience, we do suggest a couple of constraints to optimize the 3PP camera settings: (1) virtual/real body consistency of orientation facilitates the mapping of oneself into a virtual body, making it less cognitively demanding to predict the virtual body response to one's actions. That is, if the reference frame of the subject and the virtual body (as observed by a 3PP camera) are similar, it is easier for a subject to project the degrees of freedom under her control into the surrogate body. We could observe this in preliminary tests used to refine the design of experiments that follow in Chapter 4. In the occasion a lateral 3PP of the avatar was considered, which we deemed to be unnecessary in the experiment as it clearly affected performance; (2) favor a mirror metaphor if a 3PP camera is facing the user, taking advantage of our skill to control the posture of our body reflection in a mirror.

(iv)How many 3PP viewpoints should be used?

In most cases we believe that one 3PP camera might be sufficient. For instance, we might experience an expressive increase in complexity of interaction and attention dispersion if a complex behavior or direct user control are used to set the parameters of multiple cameras. However, additional 3PP cameras might be relevant if the 3PP cameras are static, or if multiple orthogonal views are adopted (e.g. to increase comprehension of space), such as in previously discussed related work [Maupu et al., 2009, Wang et al., 2011].

As mentioned earlier, we have designed and implemented two interfaces combining 1PP and 3PP. The first presents 1PP and 3PP simultaneously, while the second alternates between perspectives.

3.2.3 Design 1: Simultaneous Perspective Interface

In this interface we enforce the usage of 1PP, and adopt multiple static 3PP cameras, which are presented in viewports. We explored different metaphors to insert and manipulate these viewports: augmented reality glasses (Figure 3.7a); trunk attached (Figure 3.7b); virtual tablet

(Figure 3.7c); and floating balloon (Figure 3.7d).

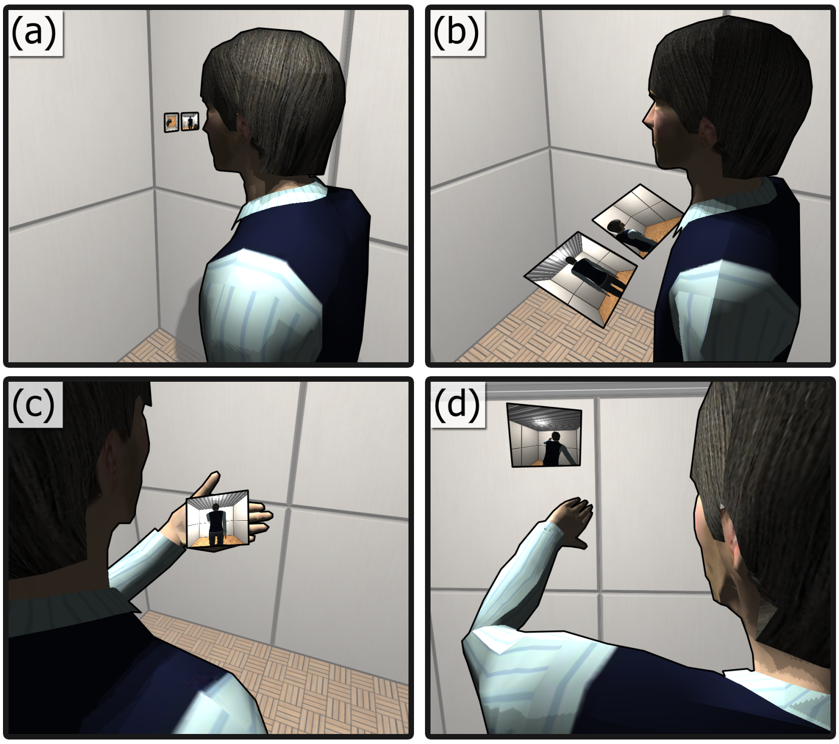


Figure 3.7 – Possible placement of viewports. If multiple viewports are preferred, one has to decide where they should be included. Here we present possible arrangements for viewports, in this figure the user takes a 1PP view of the world, which is augmented by 3PP viewports attached to specific body coordinate systems: (a) screens fixed to the head; (b) screens fixed to the chest; (c) screen fixed to the palm; (d) screen floating above the wrist. Contrast of the background was intentionally reduced to increase legibility.

The augmented reality glasses metaphor is uncomfortable when rendering with stereo due to the large disparity of objects near to user's eyes. It also partially occludes the 1PP view of the scene as it is attached to the head coordinate system. Attaching the screens to the chest coordinate system addresses these issues. However, the latter is inconvenient if the FoV is narrow, which is the case of HMDs and shutter glasses in a CAVE. The virtual tablet metaphor attaches the 3PP screen to the palm, to see the screen the palm has to be shown to the 1PP camera, what may lead to a conflict if the user wants to use that hand to interact. Thus we also propose the floating balloon metaphor, which attaches the screen to the avatar wrist and applies an offset to the vertical position of the screen, so it stands right above the wrist. The screen orientation is defined so its normal vector always points toward the 1PP camera (Figure 3.7d). Additionally, these hand based approaches may facilitate the attention switch of the user as the screens are over/near to the main interaction tools of the user, i.e. the virtual body hands. We implement a proof of concept, which is presented with a CAVE (Figure 3.8)

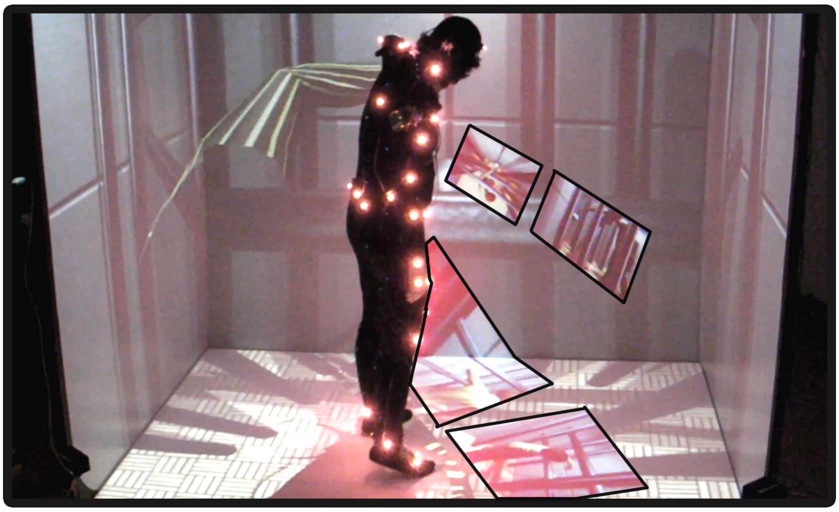


Figure 3.8 – Proof of concept of our simultaneous perspective interface. The 3PP inserts depicts two pairs of screens, one pair attached to the head and the other to the chest. These inserts are highlighted with black frames.

3.2.4 Design 2: Alternating Perspective Interface

In this interface we do not enforce a perspective in particular, the user is free to alternate between 1PP and 3PP whenever they decide to. We also adopt a single 3PP camera which is fully controlled by the user. More specifically, it rotates and translates identically to the 1PP camera, but is defined with 120 cm offset (see Figure 4.1).

The alternation of perspectives requires the implementation of a transition phase. We carefully designed one with the priority of preventing cybersickness. Cybersickness is mainly attributed to sensorial mismatch of vestibular and visual systems [LaViola Jr, 2000], i.e. when visual movement is present in the lack of its vestibular counterpart or the inability to anticipate a visual flow. To address the latter, we let the users trigger when the transition would occur, thus allowing them to anticipate the mismatch. The former is discussed below.

We then considered three different perspective transition approaches. The first followed a parametric curve trajectory with acceleration and deceleration phases. It lasted for 1 second and was intended to avoid interpenetration of the virtual body and virtual camera, known to be detrimental to the VR experience [Burns et al., 2006]. This approach was however not efficient as it required a long trajectory and continuous changes in the direction of movement (in order to dodge the virtual head). Moreover, this design often gave the false impression of real movement to the users, in preliminary tests we observed that some of them would even try to compensate for the movement, risking to lose balance. The second alternative was teleportation, it had the advantage of avoiding translation. However, teleportation is also known to cause disorientation [Bowman et al., 1997], affecting user's ability to immediately resume a task on the new point of view. Finally, we opted in favor of a very fast straight line movement. The movement only lasts for 200 ms, and the vision is slightly blurred. The short length and blurred image made it unlikely that the user could perceive the virtual camera and body interpenetration. In the experiment presented in Section 4.3, this transition allowed subjects to quickly resume their actions in the new point of view, and none of them reported feeling sickness with the fast transition. Figure 3.9 presents screenshots of this transition.

We also found support to this implementation in [Lopez et al., 2014]. There the authors compared different camera movement metaphors in order to exchange the avatar being controlled. They found that subjects generally prefer simpler straight movements or teleportation in the transition, rejecting complex curved movements.

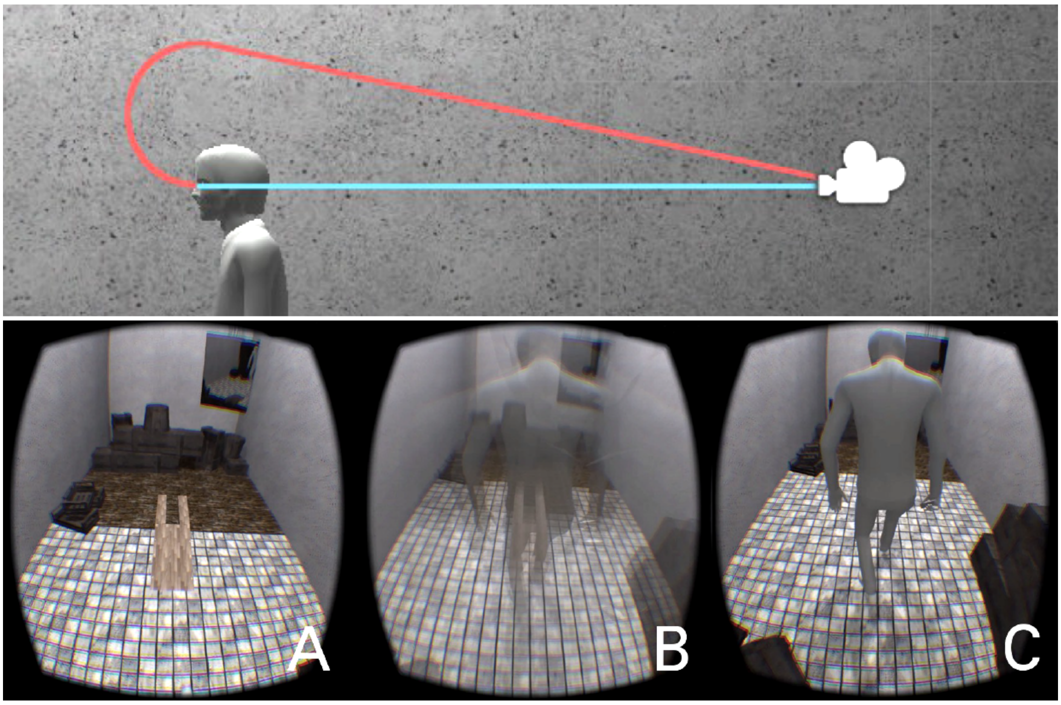


Figure 3.9 – Screenshots of perspective alternation. If the alternation between 1PP and 3PP is preferred, one has to decide how the transition will occur, and how it gets triggered. In the top, the red line shows a trajectory that avoid interpenetration, which was the first we have tried, the blue line shows the trajectory we preferred. In the bottom, a very quick and straight movement (b) is used to move from 1PP (a), to 3PP (c). This method has been tested in Section 4.3.

3.2.5 Discussion

In this section we have proposed a design space for interfaces seeking to combine 1PP and 3PP in a single visualization pipeline. To simplify its usage, we defined it in 4 design decisions: modality (simultaneous or in alternation); priority (1PP, 3PP, or not enforced); number of 3PP cameras; 3PP cameras behavior. In Table 3.1 we list the designs we implemented with those we found in literature according to the design space we propose.

Table 3.1 – Classification according to the proposed design space. (* originally a desktop interface / not a full-body interface.)

interface	modality	priority	# of 3PP cameras	3PP behavior
Design 1 (Sec. 3.2.3)	simultaneous	1PP	1+	static
Design 2 (Sec. 3.2.4)	alternating	not enforced	1	head linked
[Maupu et al., 2009]	alternating	not enforced	2 orthogonal	static
[Salamin et al., 2008]	simultaneous	3PP	1	trunk attached
[Hirose et al., 2006]*	simultaneous	1PP	unlimited	static
[Wang et al., 2011]*	simultaneous	1PP	2	cursor attached
[Stoakley et al., 1995]*	simultaneous	1PP	1	user controlled

We believe that the combination of perspectives may be applied for learning and evaluation of body postures, such as dance steps, martial arts, physiotherapeutic exercises, and preparatory training for dangerous situations. Further adding to research on physical action learning with VR, such as the study by [Bailenson et al., 2008] which shows subjects' preference for using stereoscopic rendering and a 3PP representation of the user and instructor (prerecorded animation) side by side. The non VR condition used a monoscopic view of the instructor alone (i.e. video lesson).

Although we perform no formal experiment to compare our interface design propositions, it became clear that the attention switch required in the simultaneous setup was time consuming, and that the learning curve was steeper than the alternating design. On the other hand, the alternating design presents less visual load to the subject, provide control of the 3PP camera without requiring much effort from the user, and might be easier to adapt to different tasks and applications. Thus, we would generally recommend the latter.

3.3 Synthesis and conclusion

In this chapter we explored two approaches to provide users with increased awareness of a controlled virtual body:

In the first approach we implement non-planar projections as a means to increase the FOV. We performed a short experiment where subjects wore a HMD and were instructed to perform a selection and docking task while using either Perspective ($\approx 106^\circ$ vertical FoV), Hammer or Equirectangular ($\approx 180^\circ$ vertical FoV for both) projection. Our results demonstrate some advantages, such as the reduction in search time. Other aspects, such as the simulation sickness and the sense of embodiment questionnaires seem not to be affected by the change in projection method. On the other hand, quality of interaction seems to be consistently lower, with increased time to dock and reduced docking precision. Finally, the spatial understanding of the virtual body/environment relation seems impaired as all but one subject presented an increase in the mean collisions per trial with the non-planar projections.

In the second approach we propose a design space for combining 1PP and 3PP in a single visualization pipeline. By combining perspectives we expect to be able to take advantage of their individual strengths. Namely, 1PP of a virtual body is known to consistently induce the sense of ownership of the surrogate body, while 3PP can provide constant feedback of the virtual body posture and its relation with the environment. We explored a couple of setups based in this design space, from which we chose a design that allows the subject to alternate between 1PP to 3PP at will.

Given that the performance gain of an increased FoV was limited, we believe that non-planar projections could be used for quick inspection of the environment. In such scenario, the user would switch from perspective to a non-planar projection and back as an alternative to rotating the head. This could allow for the efficient gathering of the structural information of the environment, as demonstrated in our time to reach results, and would be much like the alternating perspective interface that we propose in the second part of this chapter.

Taken together, we still expect the 1PP/3PP alternating perspective solution to fit better to our needs. It is simple to use, and provides more information about virtual body/environment than the increase in FoV offered by the non-planar projection. Thus we test that approach in terms of the sense of embodiment in Chapter 4.

Regarding the display modality, we decided to use an HMD instead of a CAVE for the remaining of this work. Indeed, the HMD is convenient as it occludes the real body from the user's field of view. Analogously to the rubber hand illusion, in which the illusion is not effective if the real hand is not occluded, seeing the own body while in a CAVE might anchor the perception of the self to the real body. Additionally, the movement distortion experiments presented in Chapter 5 can only be effective if the subject is not visually aware of their own body posture.

4 Perspective Taking and Embodiment

In this chapter we present two experiments, both of which explore factors that influence the sense of embodiment over a virtual body. In particular, we manipulate the point of view, confronting first and third person perspectives (1PP and 3PP), and visuomotor contingencies that relate the body of the subject to the virtual body.

Literature with respect to the sense of ownership and self-location of a virtual body seen from 3PP shows seemingly discordant voices. On the one hand, a series of experiments have suggested that 3PP may prevent or severely reduce the sense of ownership of a virtual body [Slater et al., 2010b, Petkova et al., 2011, Maselli and Slater, 2013, Pomés and Slater, 2013], which seems to happen as soon as the point of view ceases to overlap with the fake body [Maselli and Slater, 2014]. On the other hand, other experiments suggest a positive response [Lenggenhager et al., 2007, Lenggenhager et al., 2009, Aspell et al., 2009]. Moreover, [Noel et al., 2015] has recently shown that the full body illusion from 3PP causes alterations to the peripersonal space (i.e. the space within immediate reach of one's body parts), which drifts in the direction of the seen body.

With these experiments we expect to clarify the role of visuomotor congruency of the whole body to the sense of embodiment, with special attention to the case of 3PP.

Both experiments have a condition of high embodiment compatibility, in which the virtual body was seen from 1PP and responded to subject's movements. Alternatively, perspective could be set to 3PP, and two distinct visuomotor discrepancies have been explored. In the first experiment we introduce a 1 second delay between the performed and seen body movements, while in the second experiment we use pre-recorded motion data to animate the virtual body, thus not accounting for the movements of the subject. In the second experiment we add a passive haptics device to deliver visuo-motor-tactile congruence, and introduce a perspective condition in which the subject is able to switch between 1PP and 3PP at will. The detailed discussion of this condition has been presented in Section3.2.

4.1 Related Work

3PP is often employed in non-immersive virtual environments such as video games to increase awareness of the environment and threats to the player [Taylor, 2002], thus overcoming field of view limitations of 1PP. In VR, the usage of orthogonal third person viewpoints has been explored and was for instance recommended to help setting the posture of a motion controlled virtual body [Maupu et al., 2009]. The use of 3PP is also recommended to compensate for the compression of distance perception inherent to immersion systems such as large stereoscopic projection. This was demonstrated in a VR basketball application in which motor behavior were closer to reality in 3PP than in 1PP (speed at moment of release closer to real throw than in 1PP) [Covaci et al., 2014]. Moreover, using an HMD setup [Salamin et al., 2010] has shown that a short training is sufficient for subjects to perform distance estimation in 1PP and 3PP with similar precision. The question is therefore to know if these benefits of 3PP can be exploited without detrimental consequences to the sense of presence and the ability to embody a virtual body.

The illusory ownership of a whole body seen from outside has been demonstrated by [Lenggenhager et al., 2007]. In that experiment, the synchronous stroking at matching locations of the back of the subject and his/her virtual representation (video or mannequin) led subjects to feel that they were located in the . It is argued, analogously to the rubber hand illusion, that the integration of congruent multisensory information could lead to the sense of ownership of a body seen in extra-personal space.

The question of perspective has been further explored in VR. [Slater et al., 2010b] performs an experiment involving visuotactile congruence and perspective. They suggest that 1PP is a critical factor for the ownership of a body transfer illusion, contrasting previous studies that suggest visuotactile synchrony to be the critical contributory factor to the ownership illusion.

Closer to the aspects we explore here, [Pomés and Slater, 2013] presents a setup in which a virtual body is seen from behind. Congruent or incongruent visuotactile stimulation could be applied, and the subject could control the movements of the virtual body arms. No effect of visuotactile congruency has been found, and the ownership scores were generally low.

While we agree that perspective might have a significant impact to the sense of ownership, we sough to couple it with full body control, adding a new dimension to this subject. We also highlight that the animation algorithms used in [Pomés and Slater, 2013] could only animate a few degrees of freedom, thus being simpler than the ones we use here.

4.2 Experiment 1: Perspective and Visuo-motor Synchrony

This section has been adapted from [Debarba et al., 2015a].

Using an experimental paradigm based on full-body visuomotor synchronous mapping (the coherent replication of one's real body by a virtual body), we studied the effect of perspective (1PP vs. 3PP) and synchrony (movement delay) on the sense of embodiment as well as on the performance in a reaching task. We assessed subjective reports of the sense of agency, sense of body ownership and self-location, which relate to the sense of embodiment [Kilteni et al., 2012a]. This allowed studying how full body visuomotor synchrony is effective in inducing embodiment with respect to perspective, and in measuring the performance trade-offs for performing a reaching task with a 3PP viewpoint.

4.2.1 Materials and Methods

Equipment

A Phasespace Impulse X2 was used for motion capture; 10 cameras were used to track 38 markers attached to a motion capture suit and to the HMD. Data were acquired at a frequency of 60Hz. An Oculus DK1 was used to display images at a resolution of 640 x 800 pixels per eye. Its inertial sensors were used to obtain instantaneous head orientation, which was corrected for drift around the vertical axis using the attached optical markers (other axis of rotation do not drift). Figure 4.2b shows a subject wearing the suit and HMD. The integrated HMD sensors were used because they offer shorter latency than the optical markers.

Posture reconstruction: Full body motion capture was performed in real time. To account for body size variability, a calibration step was performed based on a standard posture (T-stance) that subjects were asked to perform. Lower and upper body of the virtual body were adjusted in scale, followed by arm adjustments. Finally, orientations of limbs, trunk and head were adjusted to closely match those of the subjects. A new iteration was performed if required. This calibration allowed for a close match of real and virtual bodies, and known limitations (e.g. incorrect arm and forearm proportions) were minimal thanks to the subjects' recruiting criteria. To animate the virtual body we used an in-house analytic IK implementation which reinforces co-location of tracked markers and end effectors positions [Molla and Boulic, 2013]. Fingers were not animated and were kept in a neutral pose. Discontinuities of posture reconstruction could occasionally occur during motion capture if the position of a marker became unknown – limitations of the optical tracking equipment – and we limited the visual consequence to the drop of one animation frame (less continuous movement). This occurred rarely during our experiment, only in case of very fast reaching movements (system could momentarily lose track of the marker) or self-occlusion.

Virtual environment: The subject stood over a flat plane, in which we avoided presentation of any potential spatial cues. A unique neutral and not textured virtual body was used for all subjects. Shadows from a parallel light source were projected in the ground right in front of the virtual body.

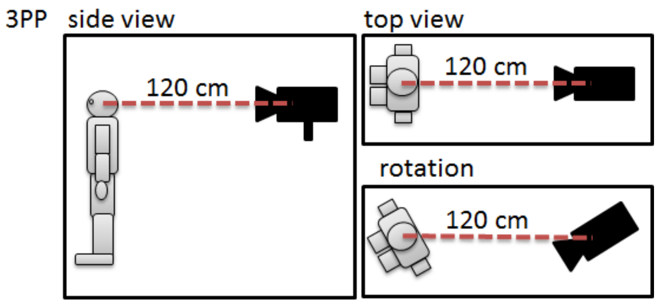


Figure 4.1 – 3PP camera behavior. The 3PP camera was set using a 120 cm offset relative to the 1PP position. The offset position was also used as its center of rotation.

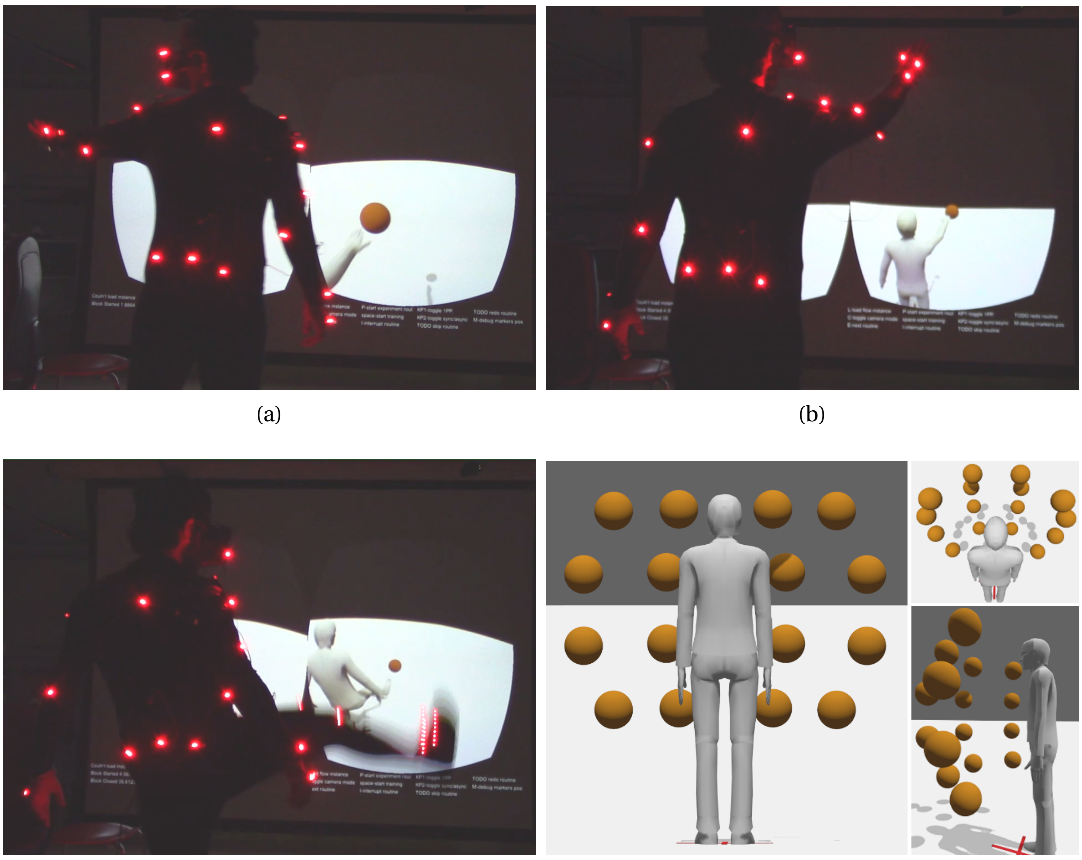
Experiment Design

For this experiment, subjects wearing motion capture sensors and HMD (Figure 4.2a-c) were asked to touch virtual targets that popped-up at a set of predefined positions within a hemisphere aligned with their virtual body (Figure 4.2d). The order of the targets was shuffled. This task was repeated for combinations of 1PP and 3PP (Figures 4.2a and 4.2b respectively) in synchronous and asynchronous visuomotor condition (1 second of delay to the motion capture data). The experiment followed a within subject factorial design with the factors perspective (1PP/3PP) and synchrony (Sync/Async). Each factor combination was repeated 3 times, for a total of 12 blocks per subject. A subject went through all the blocks of a perspective condition before switching to the other. The perspective presentation order was counterbalanced per subject, while the presentation order of the 6 blocks (3 Sync and 3 Async) within each perspective condition was randomized.

Implementation

Perspectives: 1PP – markers attached to the Oculus were used to place the virtual camera as close as possible to the subject eyes position. Rotation and translation were computed from those markers and from the inertial sensors in the Oculus; 3PP – the only difference with 1PP was an offset of 120cm backwards (behind the avatar). Translation and rotation were centered at that point (Figure 4.1). Figure 4.2ab shows how the subject sees the body and virtual environment in 1PP and 3PP respectively.

During preliminary tests we have also considered other 3PP camera settings, such as observing the virtual body from a lateral point of view. However, this conception clearly affected the perception of space, and consequently the reaching task performance. Moreover, further evidence on cognitive neuroscience suggests that misalignment of orientation affects the sense of ownership of the virtual body. [Blanke et al., 2015] proposes that the orientation match of real and surrogate bodies is a constraint to successfully manipulate bodily self-consciousness. The rationale being that the incongruent orientation results in proprioceptive and vestibular discrepancies, thus reducing the multisensory channels that drive this illusion.



(a) (b)

(c) (d)

Figure 4.2 – Overview of perspective conditions and setup. (a-c) The subject was equipped with an HMD and a motion capture suit containing 37 LED markers. (ab) Illustrates a subject performing the reaching task in 1PP and 3PP. (c) Targets could be acquired with any body part. (d) Targets were arranged on 16 pre-defined positions laying in the surface of a sphere, they were presented in a randomly shuffled order. The image projected in the screen (a-c) shows what the user is currently looking at and only serve illustrative purposes.

Synchrony: Sync – motion capture and camera translation and rotation in real time; Async – 1 second delay of motion capture and cameras translation, camera rotations were kept in real time to prevent cybersickness.

Task: Target reaching with 16 predefined positions. Targets were spread over the surface of a sphere segment of 80cm radius, centered at $x = 0$, $y = 1.2m$ and $z = 0$ (with positive y pointing up). Targets were represented as spheres of $10cm$ radius. After each reach, the subject had to return beyond a line in the floor, oriented along the lateral axis. The next target appeared 1 second after assuming the aforementioned position.

Table 4.1 – Questionnaire and results. Answers were given in a visual analog scale (VAS) ranging from 0 to 100. Median, interquartile (IQR) and p-values (Wilcoxon Signed-Rank test) for synchrony and perspective factors are presented. Results per groups (perspective presentation order: 1PP-3PP and 3PP-1PP) are also shown.

Questions			Asynchrounous		Synchronous			1PP		3PP		
During the last session there were times when...		Group	Median	IQR	Median	IQR	p	Median	IQR	Median	IQR	p
Q1. ...it felt like I was in control of the body I was seeing	(Agency)	**All**	39.5	40.4	95.9	8.4	**0.000** *	85.5	52.6	69.0	58.6	0.221
		1->3PP	53.0	56.4	94.8	7.6	**0.001** *	89.0	34.9	92.3	53.6	0.860
	Agc	**3->1PP**	37.5	28.7	97.9	9.6	**0.000** *	78.6	59.2	71.3	58.6	0.148
Q2. ...it felt that the virtual body was my own body	(Body Ownership)	**All**	45.6	40.4	89.8	19.9	**0.000** *	76.9	44.5	65.9	52.2	0.914
		1->3PP	55.2	67.4	85.6	29.7	**0.000** *	76.9	32.4	65.6	55.1	0.375
	BOwn	**3->1PP**	40.5	28.6	92.9	10.8	**0.000** *	75.8	54.3	67.0	45.0	0.222
Q3. ...it felt as if my body was located where I saw the virtual body to be	(Self-location)	**All**	50.0	34.0	91.6	22.9	**0.000** *	82.8	40.1	66.7	42.2	0.883
		1->3PP	54.1	49.0	88.7	25.6	**0.003** *	83.8	30.8	63.6	46.2	0.782
	SLoc	**3->1PP**	44.4	24.2	92.8	17.7	**0.000** *	78.6	49.4	67.9	40.8	0.755
Q4. ...it felt as if I had more than one body	(More bodies)	**All**	66.4	29.4	33.2	58.1	**0.017** *	52.0	58.4	58.4	46.0	0.197
		1->3PP	73.0	38.9	40.9	59.2	0.403	60.5	52.8	72.9	54.7	0.433
		3->1PP	62.3	18.7	23.9	32.1	**0.030** *	36.6	56.1	46.9	33.4	0.193

Procedure

Each block consisted of 90 seconds of VR exposure during which subjects were standing and performed several reaching movements using any part of their body. After each block, subjects were guided to a desktop computer and were asked to remove the HMD and to fill-in the questionnaire using a regular mouse. Questions were presented in white over a black background. The subjects were allowed to sit and rest between blocks, and were informed about how many blocks were left.

The questionnaire was adapted from [Lenggenhager et al., 2007, Tsakiris et al., 2010] to estimate subjective sense of agency (Q1), sense of body ownership (Q2) and self-location (Q3). Q4 asked whether the subject felt to have two bodies (control). A Visual Analogue Scale (VAS) was used to record answers ranging from “disagree” to “agree” (100%). The questionnaire is presented in Table 4.1.

Recruiting

In order to minimize variations on the motion capture stability and visual experience over subjects, we established four recruitment criteria: male gender; ability to focus on infinity without glasses or using corrective lenses bundled with Oculus HMD; height between 170 cm and 185 cm; body mass index (BMI) between 18 and 23.

A total of 16 subjects participated to this study (ages from 21 to 31, mean of 26, all right handed). They were recruited through an online call in the university and were paid 20 CHF per hour of their time. The experiment took between 70 to 110 minutes, depending on the setup time (calibration and adjustments). This experiment was approved by the Commission cantonale d’éthique de la recherche sur l’être humain in Vaud, Switzerland. Subjects signed a consent form and were paid 20 CHF/hour for their participation.

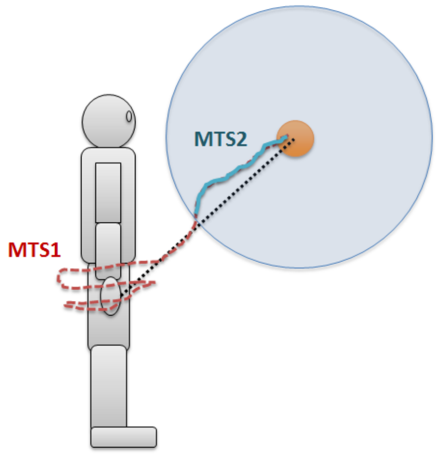


Figure 4.3 – Criteria defining the two stages of a reaching. MTS2 is measured from the last cross of 50% of the distance between initial position of the limb and target position, MTS1 is computed as MT - MTS2. The black dotted line represents the initial distance between used end effector and target. The curved line represents a hypothetical reaching trajectory.

Dependent Variables

The dependent variables are; reported sense of *Agency* (Agc), sense of *Body Ownership* (BOwn), *Self-Location* (SLoc), *Mean Time* to reach (MT), and *End Effector* choice (EE). Agc, BOwn and SLoc were measured through questions Q1, Q2 and Q3 respectively.

MT represents the mean time each subject took to reach a specific target out of the 16 possible targets (Figure 4.2a), and is relevant only for the synchronous condition. For analysis, it is split in two stages, the early one accounts mostly for visual search and movement initiation (MTS1) and the late one accounts for movement completion and target hit (MTS2). MTS2 is measured from the last cross of 50% of the distance between initial position of the limb and target position, MTS1 is computed as MT - MTS2 (4.3). As the typical velocity profile of a reaching movement tends to be symmetrical and to resemble a normal distribution [Sciutti et al., 2012], this splitting criteria allows dividing the movement where it is more likely to reach its maximum velocity. This is also where it is less likely that information expected to be part of MTS1 would affect MTS2, and vice versa.

Finally, the end effector preference ratio (EE) describes the preferred limb used for selection per given target. It is measured as the proportion of times the subject used a given limb over the total number of reaching movements he performed for that target.

Hypotheses

We identified five hypothesis that our experimental manipulation allows to investigate.

H1: The synchrony of avatar movement influences the sense of embodiment and each of its component : Agc (H1.1), BOwn (H1.2) and SLoc (H1.3). This indicates whether full body visuomotor synchrony is a factor influencing the components of embodiment.

H2: The perspective factor influences the sense of embodiment and each of its component : Agc (H2.1), BOwn (H2.2) and SLoc (H2.3). This indicates whether perspective is a factor influencing the components of embodiment.

H3: Agency is an enabling factor for body ownership and self-location. Observing a correlation between the reported Agc with BOwn (H3.1) and SLoc (H3.2) indicates whether the sense of agency for the avatar's movement is linked to other components of embodiment.

H4: Time to reach targets is influenced by perspective. In particular, we expect that 3PP presents shorter MT by reducing the required visual search time (H4.1). More specifically, we expect MTS1 to be smaller for 3PP (H4.2), and MTS2 to be equivalent across perspective change (H4.3). This analysis will highlight locations of targets showing an advantage for 3PP, or conversely. We do not consider the Async trials in this question as its effect is inherently negative.

H5: Subjects can accomplish the task in a similar manner across perspective conditions. This is assessed through variations of preferred end effector (EE) per target position across perspective.

Analysis

Questionnaire analysis for H1 and H2 was carried out using Wilcoxon Signed-Ranks Test, for which data were paired per subject across the combination of conditions (1PP/3PP and Sync/Async). Spearman correlation was computed between variables for H3.

Only the data in the Sync condition was used for performance analysis of MT (H4). Our analysis includes time to reach responses (MT, MTS1 and MTS2), and was carried out with paired samples t-test. More specifically, we compare difference on MT, MTS1 and MTS2 between perspectives for each of the 16 possible targets. Outliers were defined based on the interquartile distance (i.e. less than quartile of $25\% - 1.5 * IQR$ or greater than quartile of $75\% + 1.5 * IQR$) per perspective × target position combination. Targets that were selected using the trunk or the head were very few, and were removed to prevent bias. Therefore, from the total of 1760 trials, 1627 were kept for analysis. If a data point for a given perspective and target position combination was missing for a subject (i.e. no selection performed for that specific combination), its pair was also removed from the analysis.

The same subset of trials was used for EE choice (H5). One-sided paired t-test was used to compare perspectives for superior and inferior limbs, as well as left and right sides of the body for each target position.

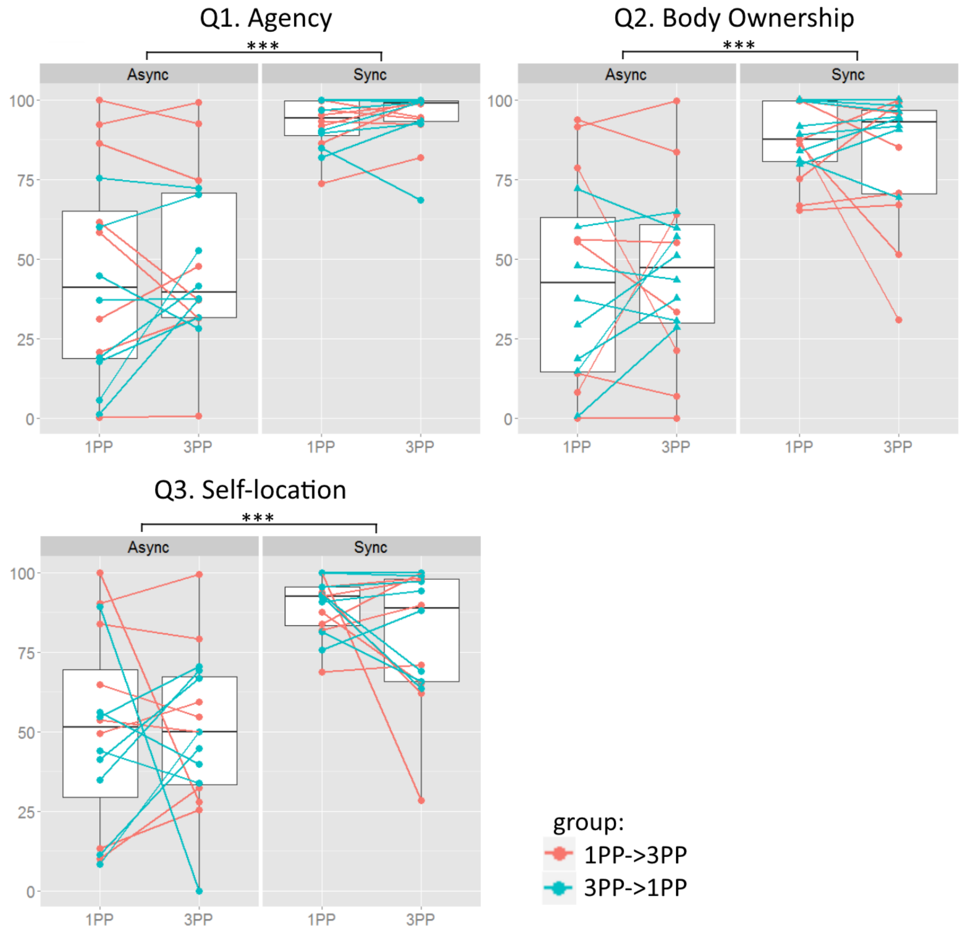


Figure 4.4 – Boxplots of reported sense of agency (Q1), sense of body ownership (Q2) and self-location (Q3). Lines represent per subject change in response across perspective; *** means significance with p<0.001

4.2.2 Results

Differences for synchrony are significant for Agc, BOwn and SLoc (all $p<0.001$, Figure 4.4), confirming H1.1, H1.2 and H1.3. Although Q4 was meant to be a control question, it also presents a significant difference ($p<0.02$), but much weaker. Perspective only has significant effect for Q4 when considering Sync condition alone ($p<0.03$). Differences for perspective are not significant for Agc, BOwn nor SLoc (Figure 4.4), thus failing to reject equality for H2.1, H2.2 and H2.3. Agc responses are positively correlated with BOwn and SLoc, supporting H3.1 and H3.2 (all $p<0.01$). Table 4.1 shows the Median, interquartile and p-values for the whole sample analysis, as well as for groups (2 groups, those who started with 1PP and those who started with 3PP). The responses to all questions except Q4 agree when statistics for unbalanced groups are taken, which demonstrates no order effect for relevant measurements.

MT shows no global advantage for any specific perspective, but selection of targets surround-

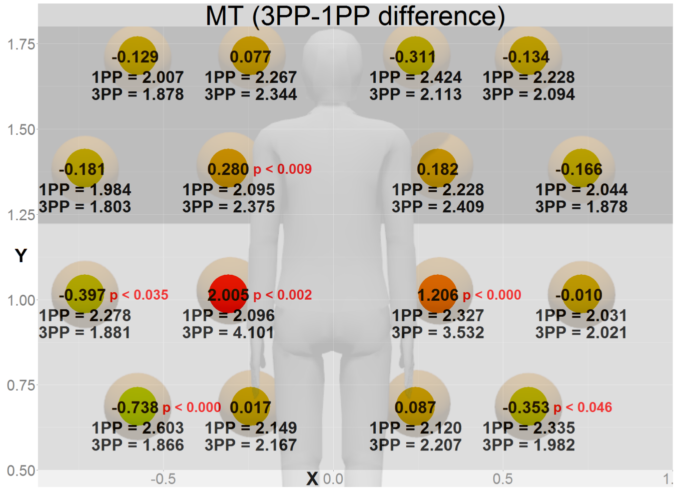


Figure 4.5 – Mean difference for Mean Time (MT) response. The redder, the bigger the advantage of 1PP; the greener, the bigger the advantage of 3PP; p-values in red indicate significant differences in a two-sided paired t-test for that specific target

ing the avatar show an advantage for 3PP, and selection of targets that may be occluded by the virtual body show an advantage for 1PP. Only 6 out of the 16 (37%) target positions presented statistically significant difference of MT between conditions (Figure 4.5). MTS1 follows the tendency of MT, with significant differences for 10 out of 16 target positions (62%). MTS1 reveals a clearer advantage of 3PP for targets that are not subject to occlusion. On the other hand, targets that are likely to be occluded presented the biggest differences of MT and MTS1. H4.1 and H4.2 are supported when visual occlusion is unlikely, but rejected otherwise. Nonetheless, MTS2 shows a clear disadvantage for 3PP as 8 out of 16 target positions presented statistically significant differences supporting 1PP (50%). In addition, only 2 out of 16 targets presented an advantage for the mean of 3PP as compared to 1PP in MTS2. Thus, providing evidence to reject H4.3. Figure 4.6 reports MTS1, MTS2 and their differences for each perspective and target combination.

Proportions of end effectors (EE) used for reaching are shown in Figure 4.7. They are similar when considering distribution for superior and inferior limbs, the difference being significant for only one target. However, lateralization seems to be unbalanced in 1PP, presumably due to handedness (subjects were all right-handed). This lateral distribution asymmetry does not seem to be present in 3PP. Laterality was significantly different for three targets located at the lower left of the virtual body. Thus, H5 is confirmed when considering superior and inferior limbs, but not entirely when considering right/left sides.

4.2.3 Discussion and Conclusion

The subjective reports of all three components of embodiment show a significant impact of visuomotor synchrony on body ownership and self-location. This is in line with experiments

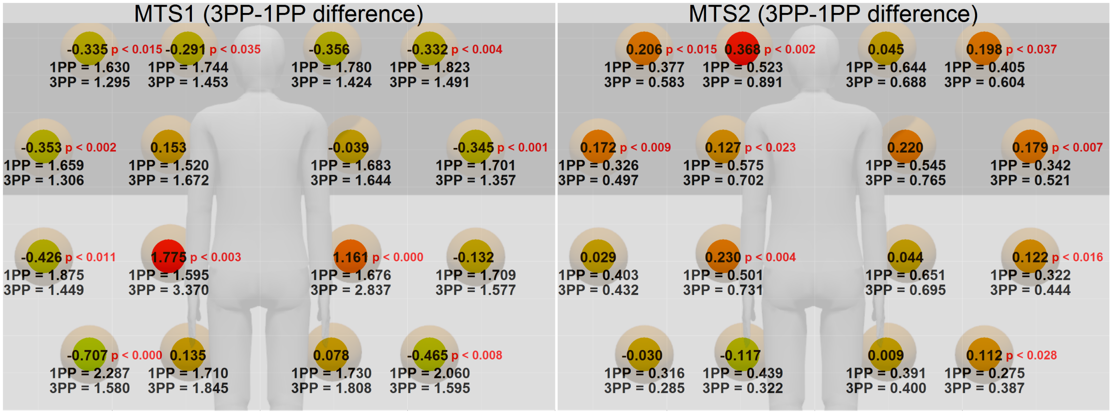


Figure 4.6 – Mean difference for Mean Time Stage 1 and 2 (MTS1/MTS2) responses. The redder, the bigger the advantage of 1PP; the greener, the bigger the advantage of 3PP; p-values in red indicate significant difference in a two-sided paired t-test for that specific target.

using full-body visuomotor synchrony in virtual mirror paradigms that elicit a high sense of ownership [González-Franco et al., 2010]. Other experiments comparing the influence of visuomotor and visuotactile congruency on body ownership also demonstrated a strong influence of visuomotor synchrony on multiple measures of embodiment [Kokkinara and Slater, 2014].

Data also show that, in the context of our full-body interaction and reaching task, perspective (1PP vs. 3PP) did not influence the subjective evaluations of embodiment. This contrasts with some previous work where the perspective change was observed to influence ownership of a virtual body [Slater et al., 2010b, Petkova et al., 2011, Maselli and Slater, 2013]. We suggest three possible interpretations for that. First, the difference between 1PP and 3PP could be present but our measures are not sensitive enough: the perspective effect could be compressed and no longer significant as compared to the effect size of visuomotor synchrony. A second interpretation is related to the active nature of our task, which differs from the experimental paradigms of the eariler studies. This reaching task required a high level of involvement, a sustained cognitive load, and potentially led to the mental state of *flow* [Csikszentmihalyi and Csikszentmihalyi, 1992]. This is also observed in computer games, for which different perspectives are all compatible with high engagement. Finally, we might face a ceiling effect of full-body visuomotor control as compared to the influence of perspective. Even though the sense of agency is decoupled from the sense of ownership in its neural basis [Tsakiris et al., 2010], ownership may be strongly driven by agency when shape and proprioceptive congruency are present [Walsh et al., 2011]. This is also partially supported by the positive correlation observed between the sense of agency and the reported body ownership and self-location.

Taken together, these results suggest that a 3PP can be used for immersive full-body reaching tasks and is compatible with a high level of embodiment into the virtual body. Only the differences in reaching behavior between 1PP and 3PP highlight some specific advantages

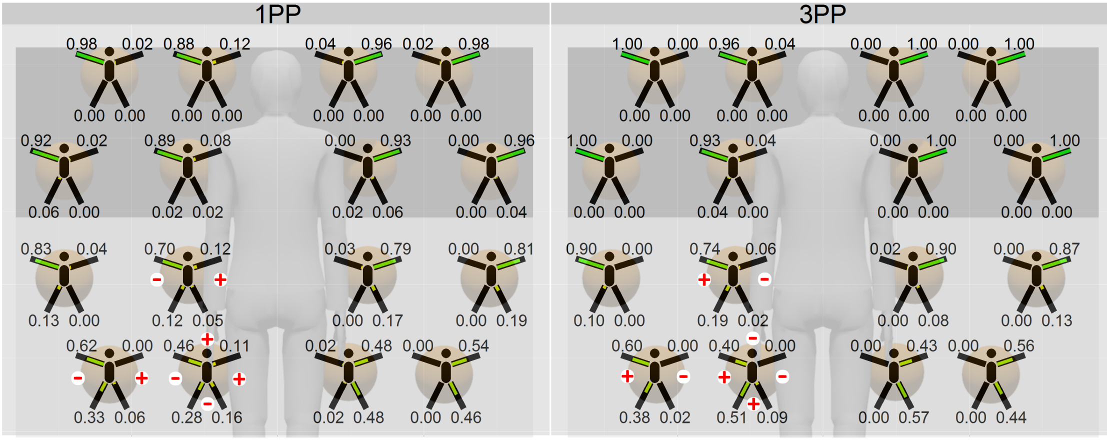


Figure 4.7 – End effector (EE) of preference for each target per perspective. +/- indicates significant difference (increase/decrease) for inferior/superior limbs and/or left/right sides of the body (e.g. + at the left of the target indicates increased usage of left EEs in that perspective). In 3PP laterality was enforced in contrast to handedness in 1PP.

according to the location of targets. The lower performance over the end of the movement (MTS2) for 3PP suggests a decrease in precision, most likely due to the reduction of targets angular size and depth cues (one may expect similar difference if smaller targets are used). On the other hand, the absolute difference between early and late stages of movement (MTS1 and MTS2) suggests a visual search advantage for 3PP, as long as the target is not subject to occlusion. Conversely, the comparison of the use of body parts (choosing to use upper or lower end effectors) suggests that our subjects did not change the way they interact with the VE. They often used the feet for lower targets in both 1PP and 3PP, suggesting that they were comfortable in exploring the full control over the virtual body independently of perspective. Finally, as we observe a crossing of the dominant hand towards targets on the opposite side only in 1PP, we believe 3PP may be used to stimulate the use of the non-dominant hand in specific applications, such as for cognitive and clinical applications (e.g. spatial neglect rehabilitation).

Additional research would be required to disentangle the interaction between subjective reports of embodiment and observed reaching behavior. A computational model of the sense of embodiment based on observations of behavior would in theory be able to provide automatic estimates of the user's level of embodiment by analyzing movement data. But our experiment does not show any direct link or correlation that would provide an obvious solution. Similarly, more experimentation would be necessary to compare our results with reaching behavior in reality. Using the appropriate perspectives could for instance compensate for the known limitations of VR for reaching [Maupu et al., 2009], throwing [Covaci et al., 2014] or other natural interaction movements. Evaluating the transfer of skill from VR training to the real situation would in turn provide information on the benefit of providing subjects with a feedback on their body and their surrounding (such as in 3PP).

4.3 Experiment 2: Perspective and Multimodal Congruence

In this experiment we assessed the effect of congruent visuo-motor-tactile feedback (full body control and passive haptic feedback) and perspective to the sense of embodiment of a virtual body. We additionally investigated how subjects behave when the possibility of alternating perspective at will is presented (see Section 3.2), and how the reported embodiment of the surrogate body in this condition compares to 1PP and 3PP alone. Our experiment consisted of a short series of tasks that the subject had to perform (or watch the virtual body performing), which ended up by exposing the subject to a virtual pit threat.

4.3.1 Materials and Methods

Equipment and Software

An Oculus development kit 2 HMD was used to display a virtual scene (960 x 1080 pixels per eye, 100 deg field of view, 75 Hz). Head tracking was performed using its inertial sensors (low latency) and corrected for drift around the vertical axis using optical tracking.

A pair of Bose® Quietcomfort 15 headphones were used for environmental noise canceling and to provide non localized white noise, thus phonically isolating the user from the real environment. Using a microphone, the experimenter could talk to the subjects directly through the headphones and provide instructions throughout the experiment.

A Wii remote was used to allow the subjects to trigger when they would like to switch the perspective in the alternating condition. The Wii controller was also been used for the mental ball drop task (detailed later in the paper). For consistency, the avatar also held an object similar to a Wii remote.

Galvanic skin response (GSR) was measured using a g.GSRsensor connected to a g.USBamp amplifier (g.tec) and recorded with the OpenViBE software [Renard et al., 2010].

A Phasespace Impulse X2 optical tracking system was used for motion capture. Our Phasespace system uses 14 cameras and 40 markers attached to a motion capture suit and to the head mounted display. A VRPN server interfaced the capture system (updated at 240 Hz) to the rendering engine (75Hz). Details for the animation of the virtual body are the same as in the experiment 1 (Section 4.2.1).

A physical object and its virtual representation were used to convey congruent visuo-tactile stimulation when walking over the pit. This manipulation is known as passive haptics, when a seen virtual object has a physical equivalent, which is calibrated to spatially match, thus rendering accurate tactile sensations. This device is made of wood and its dimensions are $140cm \times 40cm \times 10cm$. Fig 4.8a shows an overview of the experimental environment and the equipment the subject had to wear.

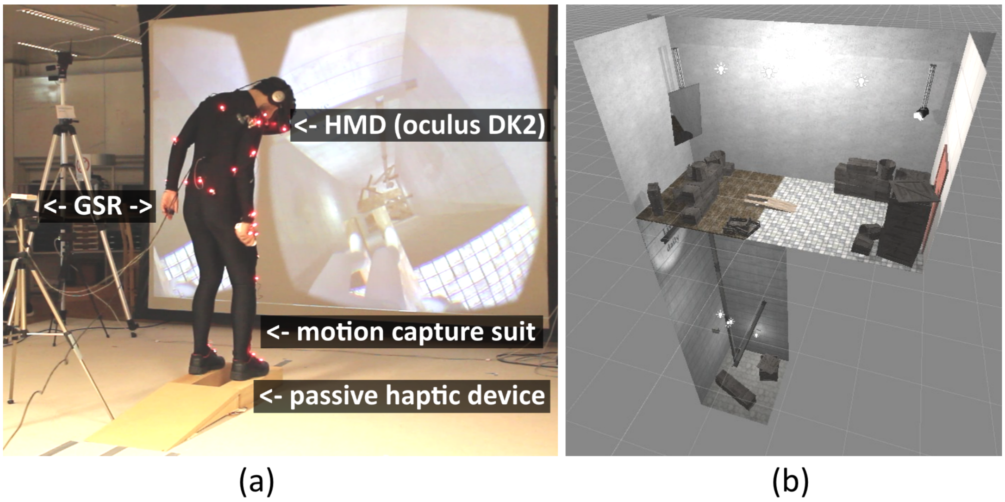


Figure 4.8 – Experiment setup and scene overview. (a) The subject was fit with a motion capture suit, an Oculus DK2, GSR sensors and a Wii remote, the image projected in the screen shows what the user is currently looking at, and only serve illustrative purposes. (b) presents an overview of the virtual scene.

The virtual environment was developed using Unity 3D, and was inspired by the pit room proposed by Meehan et al. [Meehan et al., 2002]. It featured a main room and a 10m deep virtual pit. The main room was 3.4 meter high and slightly smaller in surface than the captured space. For each session, the pit was initially covered by a wooden floor. A wooden ramp was located in the center of the scene. During a session run, the floor covering the pit would eventually fall (at the command of the experimenter), revealing the pit to the subject and leaving the virtual body standing on the wooden ramp overseeing the pit. An overview of the virtual environment is presented in Fig 4.8b.

Experimental Design

The experiment had two manipulated variables and followed a mixed factorial design, with *multimodal congruency* as the between subject variable and *perspective* as the within subject variable. Response variables were determined in order to assess components of the sense of embodiment, consisting of an embodiment questionnaire (Table 4.2), the variation of GSR following a threat event, and a mental imagery task where the subject had to estimate the time an imaginary ball would take to hit the ground (mental ball drop – MBD). The response variables are detailed later in the paper.

Subjects were assigned to one of two equal sized groups. The first group performed the experiment in a congruent visuo-motor-tactile condition (VMT group), in which subjects could control the movement of the virtual body, had to perform a sequence of tasks and could interact with a passive haptic device that stands in between the virtual body and the bottom of the pit. The second group could not control the virtual body (¬VMT group), instead subjects

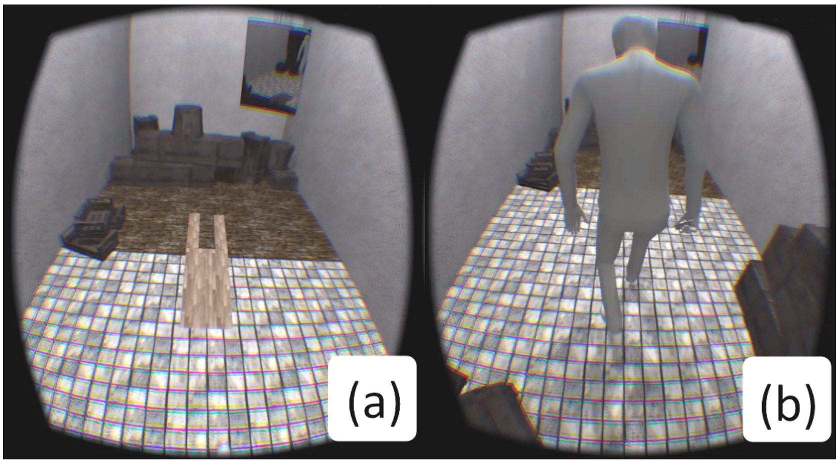


Figure 4.9 – Perspective conditions. The subject could experience the scene in three different conditions: (a) first person perspective (1PP); (b) third person perspective (3PP); or be free to alternate (ALT) between 1PP and 3PP. When in the alternate condition, subject were asked to perform at least 3 perspective switches.

had to watch the virtual body moving as recorded from one of the subjects of the VMT group. The only level of control that the ¬VMT group had was the rotations of the virtual camera. This aspect was kept across groups because it is critical to prevent cybersickness. The lack of sensorimotor feedback is expected to negatively impact the sense of embodiment of the virtual body. As the motion recordings of the VMT group were necessary for the ¬VMT condition, we ran all subjects of that group before proceeding to the second group.

Each subject repeated the experimental session three times, once for each *perspective* condition: first person perspective (*1PP*), third person perspective (*3PP*), and a novel one in which the subject could *alternate* between the 1PP and 3PP at will (*ALT*). In the *ALT* session, subjects could decide when to trigger the perspective switch by pressing a Wii remote button with the right thumb, they were also instructed to perform this action at least three times. *Perspective* presentation order was counterbalanced.

A perspective transition took 200ms, and consisted of a quick and straight movement between two endpoints (Fig 4.9). The endpoint defined by 1PP was the position in between the virtual body eyes, while the endpoint referent to 3PP had a 120cm offset toward the back of the scene (Fig 4.8) and translated relative to the 1PP endpoint (Fig 4.1). This 3PP endpoint was chosen to prevent the point of view from standing directly over the pit as it gets revealed. That is, in 3PP the virtual body would get exposed to the pit, while the visual point of view would remain over a safe area (the concrete floor. More details on the implementation of the ALT *perspective* condition are presented in Section 3.2.4.

Procedure

After reading the information sheet and completing the informed consent form, subjects were asked to fill in a characterization form with questions about their background (other experiments, experience with HMDs ...) and physical characteristics (height, weight and age).

Then the experimenter played a video demonstrating the stages of a session and subjects were asked to wear the motion capture suit. Subjects in the *VMT* group had to undergo the motion capture calibration at this point. A brief training on how the mental ball drop (MBD) task should be performed followed, using the laboratory floor as a reference.

Finally, the experimenter helped the subject fit the HMD and the noise canceling headphones, and tested the verbal communication through microphone. The GSR electrodes were placed in the left hand and the wii remote in the right hand. The subject then went through an experiment session. After the session was complete, the image on the HMD went black, and instructions of the MBD task appeared. The task was repeated 5 times, and then the experimenter removed the HMD and the headphones and asked the subject to fill in the embodiment questionnaire (Table 4.2). The session procedure was repeated three times, once per *perspective* condition.

After the experiment subjects filled-in a post experiment questionnaire about their perspective of preference for different stages of the session, as well as whether they considered the floor of the laboratory or the floor of the virtual environment during the MBD task.

This experiment was approved by the Commission cantonale d'éthique de la recherche sur l'être humain in Vaud, Switzerland. Subjects signed a consent form and were paid 20 CHF/hour for their participation.

Session Overview

An experimental session was divided into 4 stages: REACH, WALK, WAIT and OBSERVE. For the VMT group the session started with a short communication to check the setup, then:

REACH: the subject had to reach 12 targets appearing around him/her (Fig 4.10a). There were six ground and six air-targets activated one after the other in a shuffled order, and between each target reach the subject had to place back both feet on a central target. The targets were placed such that they were at equal distance to the central target (ground targets), and to the chest of the participants (air targets).

WALK: a 13th target eventually lights up in front of the wooden ramp, inviting the subject to walk from the initial position to the edge of the ramp, i.e. the passive haptic device (Fig 4.10b). The central target and the front of the ramp were separated by 2.1 meters.

WAIT: once the subjects arrive to the end of the ramp, they were orally instructed – through their headphones – to feel the edges of the ramp with their feet, sensing the passive haptic device while observing the virtual body simultaneously touching it (Fig 4.10c). During this event the experimenter would press a button, and the floor would fall down within 1 to 5 seconds (random), with a cracking sound (Fig 4.10d).

OBSERVE: the floor fall event marked the transition to the OBSERVE stage. In this stage the subjects were asked to read some words in the pit wall opposite to where the virtual body stands, so that they had to face the pit.

For the ¬*VMT* group the virtual body was driven by the data recorded from the *VMT* group. No passive haptic device was used and the subject did not have to act to complete the session. The subject was told that the virtual body would move by itself, and that (s)he should pay attention to what the virtual body was doing. The camera position also moved according to the recording, but the camera rotation could still be controlled by the subject. We kept this level of control due to its critical role preventing cybersickness [LaViola Jr, 2000]. The session started with a short communication, and further communication followed to remind subjects to pay attention to the virtual body, and that they could not control it (in case they tried to). To assign the recordings to subjects in the ¬VMT group we have paired VMT and ¬VMT subjects, the pairing was random and assured that the subjects in both groups were assigned to the same perspective order, i.e. a ¬VMT subject that did the experiment in the 1PP, 3PP and ALT order used the recording of a VMT subject who did the experiment in that same order. We had to repeat some of the VMT group recordings due to a technical issues with the recording software used for the first 5 subjects.

Response Variables

The *questionnaire* was designed to assess the senses of agency, ownership, self-location and the reaction to threat. It contained 10 questions, two for each measure, plus two controls. Questions were formulated based on related experimental protocols [Longo et al., 2008, Caspar et al., 2015, Lenggenhager et al., 2007], and designed for 7-point likert scale answers, ranging from “Strongly DISAGREE” (-3) to “Strongly AGREE” (+3). We use the mean of the two related questions as the value to four response variables; *Ownership*, *Agency*, *Self-location* and *Threat*. The questions were presented in a random order after each session.

Galvanic Skin Response: GSR was recorded to assess physiological responses to threat (floor fall event). The threat is expected to increase arousal, affecting skin conductance. We expect a GSR increase due to the threat, and we expect that the increase magnitude will be related with sense of ownership. This type of measurement has been shown to be valid in stressful virtual environments by Meehan et al. [Meehan et al., 2002], being responsive even after multiple sessions with the same subject. On the other hand, GSR tends to present high inter-subject variability, making it less reliable for confronting VMT and ¬VMT groups.

The electrodes were placed on the index and little fingers of the subject and the GSR was recorded at a sampling rate of 512 observations per second. Our GSR response variable is

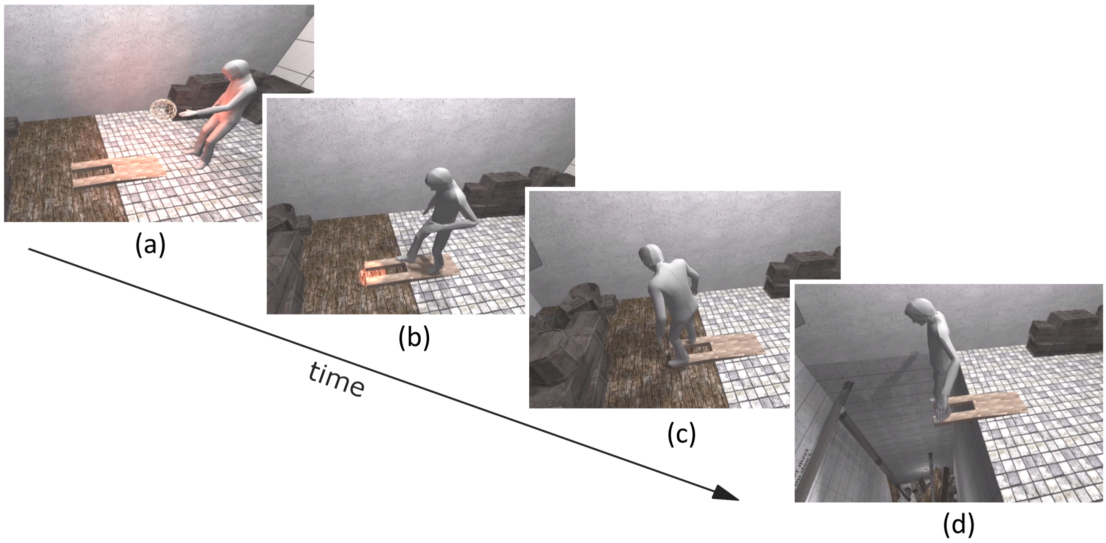


Figure 4.10 – Overview of the session stages. (a) First the subject has to reach for targets that can appear either in the air or in the floor (REACH stage); (b) a final target invites the subject to walk to the wood platform (WALK); (c) once on the platform, the subject is asked to feel the edges with their feet (WAIT); (d) finally, the wooden floor beneath the platform collapses, revealing the pit to the subject (OBSERVE). Subjects in the ¬VMT group do not perform these task, instead they watch recordings from the VMT group. The session was followed by the mental ball drop (MBD) task and an embodiment questionnaire.

defined as the difference between the median GSR in the interval between 1 and 6 seconds following the floor fall event, minus the median GSR in the 5 seconds preceding this event. Median GSR was preferred because some subjects presented a response that could go beyond the $\approx 6\mu S$ (microsiemens) recording window that our setup allowed. By using the $median$ instead of $mean$ these subjects could be kept in the analysis.

Mental Ball Drop: MBD is a mental imagery task adapted from [Blanke and Metzinger, 2009]. In this task, the subjects estimate the time a ball would take to fall down from their hand to the floor. This measurement was performed at the end of each session, when the virtual body was standing on the wooden ramp at the top of the pit. It is intended to assess self-location in reference to the pit in 1PP and in 3PP.

Before performing this task the screen turned black, and the measurement was then performed with the subjects unaware of their surroundings. Subjects were instructed to press and hold the trigger button of the wii remote controller to release the virtual ball, and to release the trigger button when they estimated that the ball have reached the floor. Subjects were not instructed about which floor they should consider (lab floor, point-of-view floor or pit floor). It was repeated five times for each session. The median of these five trials gives the MBD time estimation for a given subject and condition.

MBD is meant to detect whether the subject have similar time estimation in 1PP and 3PP.

Table 4.2 – Embodiment questionnaire applied in the end of each session. Answer was given in a 7 point likert scale ranging from strongly disagree (-3) to strongly agree (3). The variable corresponds to the mean answer to both questions.

Response variable	Question: During the last session ...
Agency	Q1 ... it felt like I was in control of the body I was seeing Q2 ... whenever I moved my body I expected the virtual body to move in the same way
Ownership	Q3 ... I felt as if I was looking to my own body Q4 ... it felt that the virtual body was my own body
Self-location	Q5 ... it felt as if my body was located where I saw the virtual body to be Q6 ... it seemed as if I were sensing the movement of my body in the location where the virtual body moved
Threat	Q7 ... I felt as if the pit posed a threat to myself Q8 ... it felt as if I could get hurt if the virtual body was to fall in the pit
More bodies	Q9 ... it felt as if I had more than one body
Turning virtual	Q10 ... it felt as if my real body was turning virtual

Consistently shorter times in 3PP could indicate weak sense of self-location, as the subject might be using the bottom of the pit in 1PP, and the floor under the camera in 3PP.

Reach performance: For the VMT condition only, we computed the median of the time to reach the targets with the hand. It is meant to assess a possible difference in performance between 1PP, 3PP and ALT conditions.

ALT usage: To assess how subjects act in the ALT condition we compute the proportion of time spent in 1PP (*p.time.1PP*). We also consider this variable for the different stages of a session (REACH, WALK, WAIT and OBSERVE). We also look for correlations between *p.time.1PP* and other response variables.

Analysis

For the response variables *agency, ownership, self-location, threat, GSR, more bodies* and *turning virtual*, the analysis was carried using mixed design analysis of variance (ANOVA) with *perspective* (1PP vs. 3PP vs. ALT) as a within subject factor, and *multimodal congruency* (VMT vs. ¬VMT) and *perspective order*(*p.order*: 1PP-3PP-ALT vs. 1PP-ALT-3PP vs. 3PP-1PP-ALT vs. 3PP-ALT-1PP vs. ALT-1PP-3PP vs. ALT-3PP-1PP) as between subject factors. We included *p.order* to verify if the comparison of *perspective* levels may have affected the response variables (i.e. does the perspective used in the prior sessions interferes with the response given for the current session?).

The reaching performance was assessed with a two-way multiple comparisons ANOVA, with

Target position (Ground vs. Air) and *Perspective* (1PP vs. 3PP vs. ALT) as the within subject variables.

As ANOVA assumes that the residuals of the model fit will follow a normal distribution, we tested this assumption with the Shapiro-Wilk test. If residuals are deemed not normal we transform the response with a Box-Cox transformation y^{λ}, which does not alter the order of the response values (monotonic transformation). We favored a λ close to 1 in order to minimally distort the data. The use of Box-Cox transformation is a common procedure to improve data distribution [Box and Cox, 1964].

We conducted post-hoc analysis with pairwise t-tests and Holm-Bonferroni correction when a significant main effect of *perspective* or interaction between *perspective* and *multimodal congruency* was found. For the latter we select a subset of possible comparisons in order to limit the correction of the alpha significance level. More specifically, we fix the value of one of the variables, and test for the combinations of the other, and vice versa. This yields a total of 9 comparisons. We do not perform any post-hoc for significant effects related to *p.order*, and simply report that a statistically significant difference has been found.

Regarding the behavior of subjects while in the ALT condition, we evaluate whether the session stage and multimodal congruency have an effect on the choice of perspective. We also look for correlations between *p.time.1PP* and other response variables.

Statistical analysis was conducted using R.

4.3.2 Results

A total of 48 subject participated on this experiment (8 females, age between 19 – 30, mean 22.6). All had normal or corrected to normal vision, normal physical and psychological condition and did not suffer from acrophobia. For technical reasons and for optimal use of the motion capture system, we also limited recruitment to subjects with height from 165 to 190 cm, and body mass index in the range from 18 to 27. Only 4 subjects reported having participated in an experiment using VR in the past, while 17 reported having tried a head mounted display (HMD) in the past, one of which with weekly frequency.

Questionnaire

The overview of questionnaire results is presented in Fig 4.11. Details of the post-hoc statistical tests are presented in table 4.5.

Agency: agency response was affected by *multimodal congruency* ($F_{1,36} = 97.7$ $p < .001$), *perspective* ($F_{2,72} = 8.7$ $p < .001$), as well as their *interaction* ($F_{2,72} = 3.4$ $p < .039$). The post-hoc of the interaction indicates a significant effect of *multimodal congruence* for all *perspective* conditions (VMT > ¬VMT). The sense of agency was significantly lower for 3PP when *multimodal*

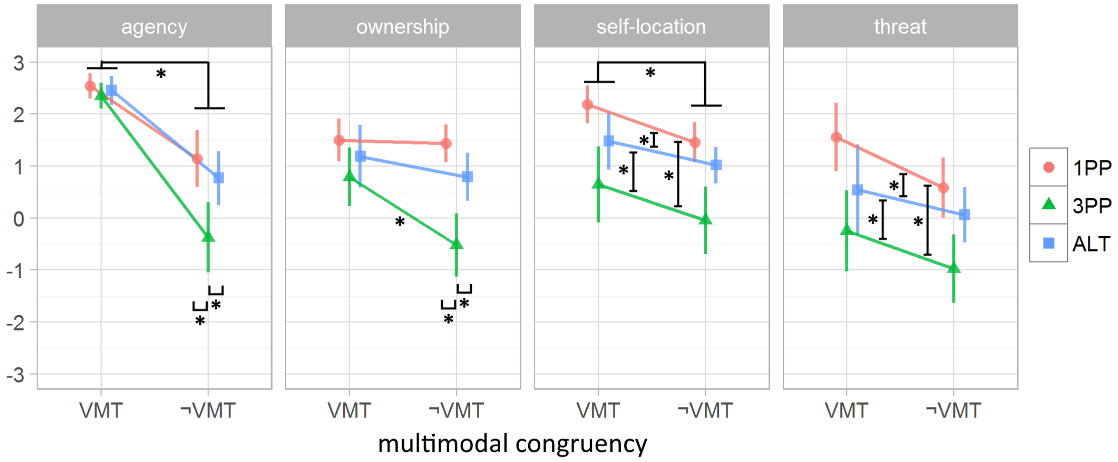


Figure 4.11 – Reported senses of agency, ownership, self-location and threat. The *perspective* factor could be set to 1PP, 3PP, or ALT (alternate between 1PP and 3PP). The *multimodal congruency* factor comprised two groups of subjects, VMT and ¬VMT (congruent and incongruent visuo-motor-tactile feedback respectively). Error bars represent the confidence interval of the mean (CI).

Table 4.3 – Results per experimental condition. Answers were given on a 7 point likert scale ranging from strongly disagree (-3) to strongly agree (3). The variable corresponds to the mean answer value given to the grouped questions.

	VMT			¬VMT		
Response	1PP	3PP	ALT	1PP	3PP	ALT
Agency	$2.54 \pm .61$	$2.35 \pm .62$	$2.46 \pm .71$	1.15 ± 1.37	$-.38 \pm 1.69$	$.77 \pm 1.29$
Ownership	1.5 ± 1.03	$.79 \pm 1.41$	1.19 ± 1.52	$1.44 \pm .91$	$-.52 \pm 1.53$	$.79 \pm 1.16$
Self-location	$2.19 \pm .91$	$.65 \pm 1.84$	1.48 ± 1.38	$1.46 \pm .98$	$-.04 \pm 1.62$	$1.02 \pm .88$
Threat	1.56 ± 1.66	$-.25 \pm 1.96$	$.54 \pm 2.19$	$.58 \pm 1.45$	$-.98 \pm 1.66$	$.06 \pm 1.33$
GSR	$.72 \pm .54$	$.46 \pm .43$	$.67 \pm .69$	$.56 \pm .69$	$.40 \pm .52$	$.53 \pm .54$

congruency was not present (1PP:¬VMT and ALT:¬VMT > 3PP:¬VMT).

Ownership: a significant main effect of *multimodal congruency* ($F_{1,36} = 4.5$ $p < .042$), *perspective* ($F_{2,72} = 22.8$ $p < .0001$) and their interaction ($F_{2,72} = 5.2$ $p < .008$) was found. Post-hoc of the interaction indicates that the response score in 3PP:¬VMT was significantly lower than 1PP:¬VMT, ALT:¬VMT and 3PP:VMT. The average ownership response was always positive when *multimodal congruency* was present, with no significant difference between *perspective* conditions in this case. It suggests that the lack of *multimodal congruency* negatively affects ownership only for 3PP.

Self-location: showed a significant effect of *multimodal congruency* (VMT > ¬VMT, $F_{1,36} = 4.3$, $p < .046$), *perspective* ($F_{2,72} = 33.8$, $p < .0001$) and of the interaction between *perspective* and *presentation order* ($F_{10,72} = 3.1$, $p < .003$). Post-hoc analysis of the perspective factor shows a significant difference between all three conditions: 1PP > 3PP and ALT, and ALT > 3PP. The interaction with *p.order* suggests that the perspective presentation order influenced the reported self-location. Specifically, subjects starting the experiment in 1PP or ALT gave lower self-location scores to 3PP, while subjects starting in 3PP gave similar scores to all *perspective* conditions (Fig 4.12).

Threat: was significantly affected by the *perspective* factor ($F_{2,72} = 21.4$ $p < .0001$). Post-hoc shows a significant difference for all *perspective* comparisons (1PP > 3PP and ALT, and ALT > 3PP). Although Fig 4.11 may suggest a consistent decrease of Threat score in the ¬VMT condition, the statistical test failed to reject the equality ($F(1,36) = 3.4$, $p > .07$).

More bodies: a significant effect of *perspective* and its interaction with *multimodal congruency* was found ($F_{2,72} = 4.3$ $p < .017$ and $F_{2,72} = 6.8$ $p < .003$ respectively). Post-hoc analysis has shown statistically significant difference with 3PP:VMT and 1PP:¬VMT > 1PP:VMT ($t_{23} = 4.56$ $p < .002$ and $t_{46} = 3.12$ $p < .03$ respectively). It suggests higher subjective agreement with this control question when ownership might be expected, despite sensory or perspective manipulation. However, the interaction between *multimodal congruency* and *presentation order* ($F_{5,36} = 3.1$ $p < .021$) also suggests that presentation order played a role on the interpretation of the question. This interaction is not detailed.

Turning virtual: a significant effect of *perspective* was found ($F_{2,72} = 16.4$ $p < .001$). Post-hoc analysis shows that 1PP and ALT > 3PP ($t_{47} = 4.83$ $p < .001$ and $t_(47) = 4.53$ $p < .001$ respectively).

Galvanic Skin Response

Eight subjects were excluded from this analysis due to missing data or to failing GSR connectors for at least one of the 3 sessions. The threat event caused a significant increase of the median for all 6 possible combinations of conditions as computed by a pairwise Wilcoxon summed-rank test. The relation between increased GSR and the threat can be visually accessed in Fig4.13 and Fig4.14. When comparing the increase observed across the the levels of *perspective* and *multisensory congruence*, ANOVA shows a significant effect of *perspective* ($F(2,76) = 6.2$, $p < .004$). Post-hoc shows a significant stronger response in 1PP and ALT as compared to 3PP ($t(39) = 3.4$ $p < .005$ and $t(39) = 2.6$ $p < .027$). The difference between 1PP and ALT was not significant ($t(39) = .94$ $p > .35$).

GSR also presented a positive correlation with the Threat question, but not with Ownership, Agency or Self-location. This suggests that the GSR was effectively related to how threatened the subject felt, validating the threat event. On the other hand, this measurement is usually expected to correlate with the sense of ownership [Petkova et al., 2011], although other experiments have failed to find such correlation [Kokkinara and Slater, 2014].

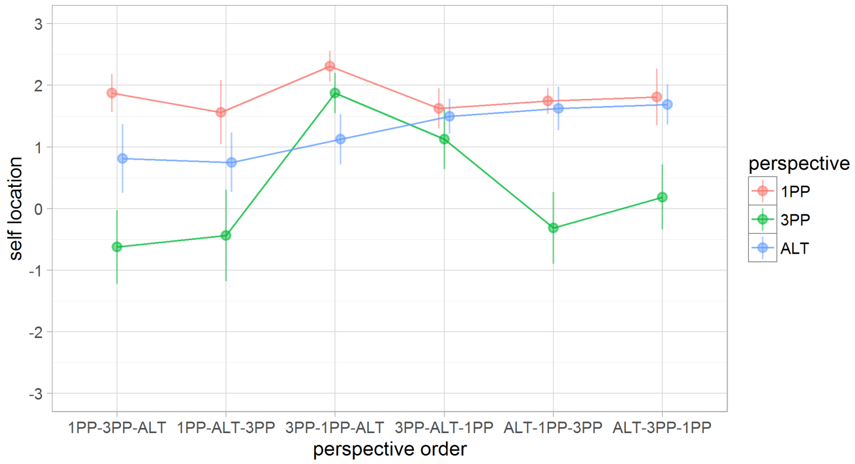


Figure 4.12 – Reported sense of *self-location* at different levels of perspective and p.order factors. Perspective order seems to influence reported sense of self-location for the 3PP condition. e.g. when answering self-location questions for 3PP after 1PP or ALT subjects tended to provide lower responses, which suggests a comparison bias.

Mental Ball Drop

The MBD time for the subjects that reported using the virtual environment floor were similar. One tailed t-test failed to reject that the time in 3PP is as high as 1PP, which makes it unlikely that subjects performed the task differently across the different *perspective* conditions. Times for these subjects in 1PP and 3PP conditions are shown in the supplementary material (4.16).

Reaching Performance

Task performance is only valid for the *VMT* group. The factor *perspective* had no significant influence on task performance (i.e. time to reach targets), ($F(2,42) = 1.59$, $p > .21$), results are shown on supplementary material (4.17).

ALT Condition Analysis

Subjects performed 2 to 30 perspective switches, with mean±SD of 11 ± 5.6. Two subjects performed less perspective changes than instructed by the experimenter. The mean±SD proportion of time spent in 1PP was $.68 \pm .13$. That is, nearly one third of the time in ALT condition was spent in 3PP. The breakdown of the proportion of time spent in 1PP during each stage of the ALT session is available in supplementary material 4.18, the graphic presents a boxplot with median and interquartile ranges to give a better picture of the group preference. Notably, overall perspective choice seems to shift to 1PP once the reaching task is complete. 1PP was especially preferred by VMT group when they had to complete the walking task. This was not the case for the ¬VMT group, who had no practical incentive to change perspective (the

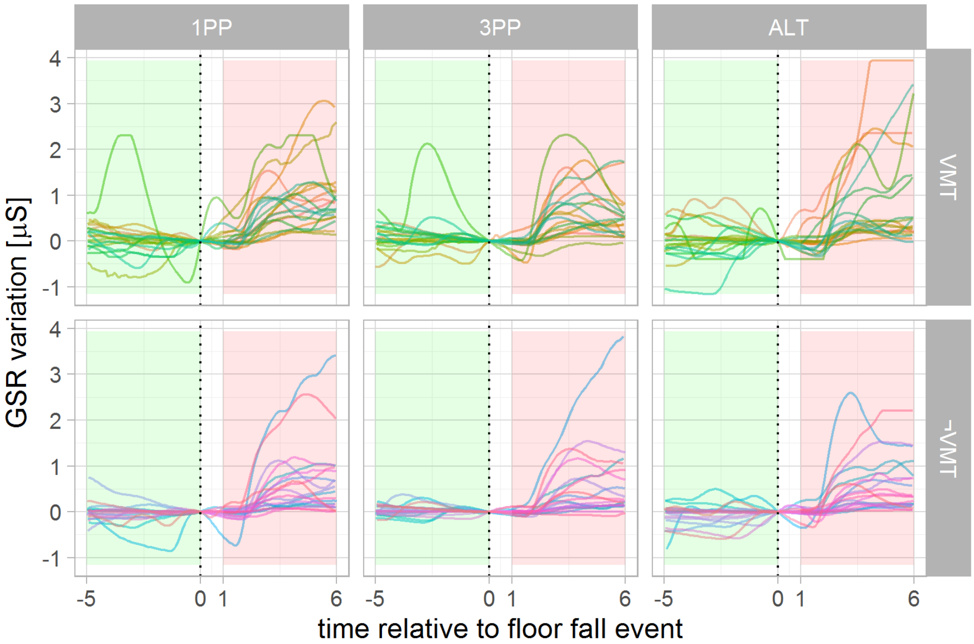


Figure 4.13 – Floor fall event locked GSR variation (units in microsiemens). The green shaded area highlights the time interval used to compute median GSR before floor fall event, while the red shaded area highlights the interval used to compute median GSR after floor fall event. Each line represents the GSR of an individual subject.

stage is completed independently of their actions). The walk stage was the only presenting a significant difference between the groups, as analyzed by a paired t-test ($t(35) = 2.88$, $p < .01$).

The proportion of time in 1PP presents a significant positive correlation with the reported sense of self-location ($r = .29$ $p < .05$) and threat ($r = .33$ $p < .022$), but do not correlate with agency ($r = -.43$ $p > .77$) and ownership ($r = .12$ $p > .4$). The latter suggests that the possibility of alternating perspective had no influence to the sense of ownership of the virtual body.

4.3.3 Discussion

In this study we manipulated visual perspective (1PP, 3PP and ALT) and multimodal congruence (VMT and ¬VMT). We assess the sense of embodiment with a questionnaire and the change in galvanic skin response due to a threat. Our threat was effective, and a clear and significant increase in GSR could be observed following the threatening event for every condition. Subjects could successfully perform all stages of all the sessions (only the VMT group had to be active).

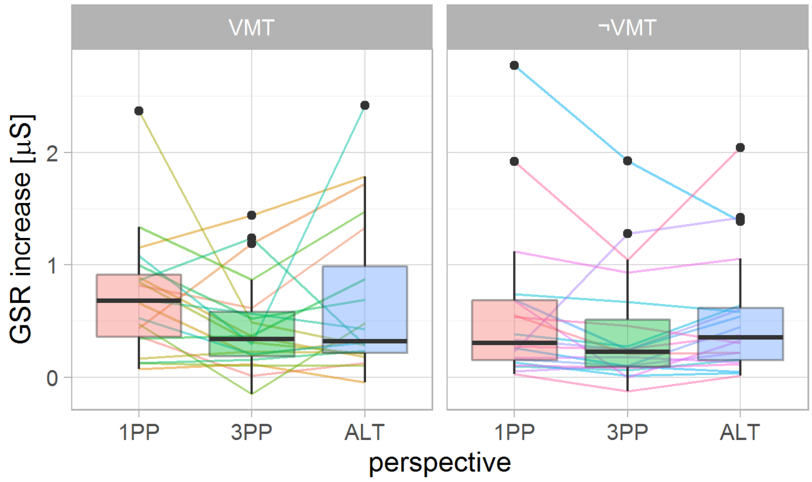


Figure 4.14 – GSR difference resulting from the floor fall event. Difference between the medians of the intervals preceding (5 to 0 seconds before) and succeeding (1 to 6 seconds after) the floor fall event was significant in all conditions. Moreover, a significant difference between 1PP and 3PP was also observed.

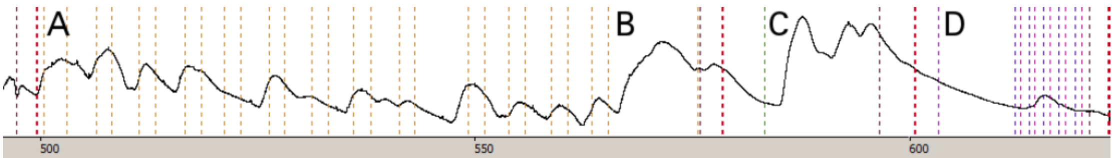


Figure 4.15 – Example of a GSR record of a session. The vertical lines indicate events transmitted by Unity to OpenVIBE. The REACH stage (A to B) tends to take most of the session time. For this specific signal one can observe anticipatory increase of arousal when approaching the region of the threat (B), and on the onset of the threat (C). (D) marks the end of this session and the start of the MBD task.

Effect of multimodal congruence

The experimental manipulation of multimodal congruence had the expected effect on the 3PP condition. The 3PP-VMT group reported a significantly stronger sense of agency, ownership and self-location than the 3PP-¬VMT group.

In 1PP the multimodal congruence effect was verified for agency and self-location, but not present for ownership. Thus suggesting a strong effect of perspective to the sense of ownership only when no other congruent sensorial clues exist. This is an appealing advantage for 1PP, as it suggests that observing the virtual body from a natural point of view while only controlling camera orientation is sufficient for the subject to self identify with the fake body, independently of proprioceptive and tactile congruence.

Moreover, even though the response to the agency questions was significantly inferior for ¬VMT, its absolute value is still positive, unveiling a degree of agreement with the sense of

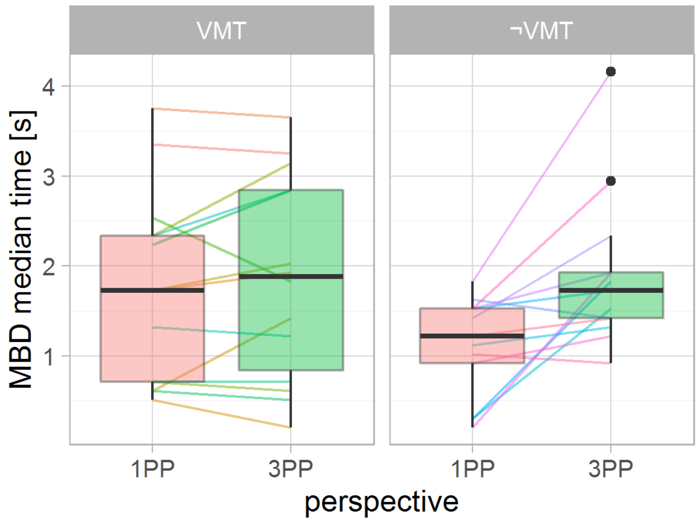


Figure 4.16 – Boxplot with median and interquartile range of MBD time.

agency question statements. These results find support on the recent work of [Kokkinara et al., 2016]. In their study, seated subjects developed the feeling of agency and ownership of a walking virtual body, but only when the externally controlled virtual body was experienced from a 1PP. The authors make the argument that, in line with the more subjective account of agency proposed by [Vosgerau and Newen, 2007, Synofzik et al., 2008b], the intention to walk may have been produced during observation, driving the self-attributing that they report. They also make a link with the findings of [Banakou and Slater, 2014], who suggests the possibility of inducing agency of an action of a virtual body – that the subject did not performed – as a result of a currently strong sense of ownership of that body. In fact, our ¬VMT condition closely replicates their experimental paradigm – with the exception that our task had higher complexity –, and our agency and ownership results are compatible with theirs, supporting their view.

1PP vs 3PP

Notably, our statistical analysis failed to reject the equivalency of ownership between 3PP and 1PP in the VMT group in questionnaire responses. Although there is a clear difference for the ¬VMT group and consistent evidence in literature that 1PP act as a decisive factor [Slater et al., 2010b, Petkova et al., 2011, Maselli and Slater, 2013, Maselli and Slater, 2014], the questionnaire results suggest that most of its influence to the sense of ownership could be mediated by multimodal congruence.

On the other hand, a significant difference was found in the GSR measurement. This measurement has been linked to the sense of ownership [Petkova et al., 2011], but its reliability as a proxy to ownership is unclear. In our study, GSR correlates with threat questions, validating the physiological measurement and questionnaire relation, but it did not correlated with the

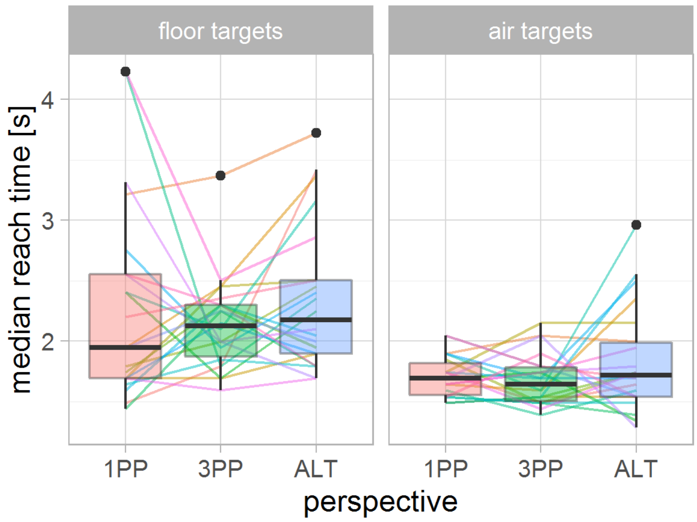


Figure 4.17 – Boxplot with median and interquartile range of time to reach floor and air targets per perspective. Performance was similar across perspective conditions. This response variable is only valid for the VMT group.

ownership.

It remains unclear if the sense of ownership in 3PP reported in the questionnaire relates to higher order processes that provide a judgment of ownership, influenced by agency and the engagement in an involving task [Synofzik et al., 2008a] (such as reaching targets), or if it is a product of the sensorimotor contingencies. A compelling new measurement that could clarify whether action in 3PP can boost the sense of ownership are those assessing the peripersonal space. Notably, [Noel et al., 2015] has shown that the classical 3PP full body illusion [Lenggenhager et al., 2007] results in a spatial drift of the peripersonal space with relation to the subjects body. More specifically, the boundaries of the peripersonal space are projected forward, toward the seen body. This protocol could help to disentangle the contribution of sensorimotor and task involvement to the sense of embodiment.

Moreover, the response to the sense of self-location for 3PP was higher for subjects who experienced this condition first, revealing the tendency to make relative judgments with respect to this response variable. This indicates that the within subject design for the Perspective variable had an impact on self-location response.

Alternating perspective

Nevertheless, the ability to choose the point of view resulted in embodiment responses that were similar to the exclusive 1PP condition. Thus, we conclude that the ALT condition is a viable alternative for VR applications to maximize the sense of embodiment, without compromising the contextual information that 3PP can provide nor the stronger bound to the

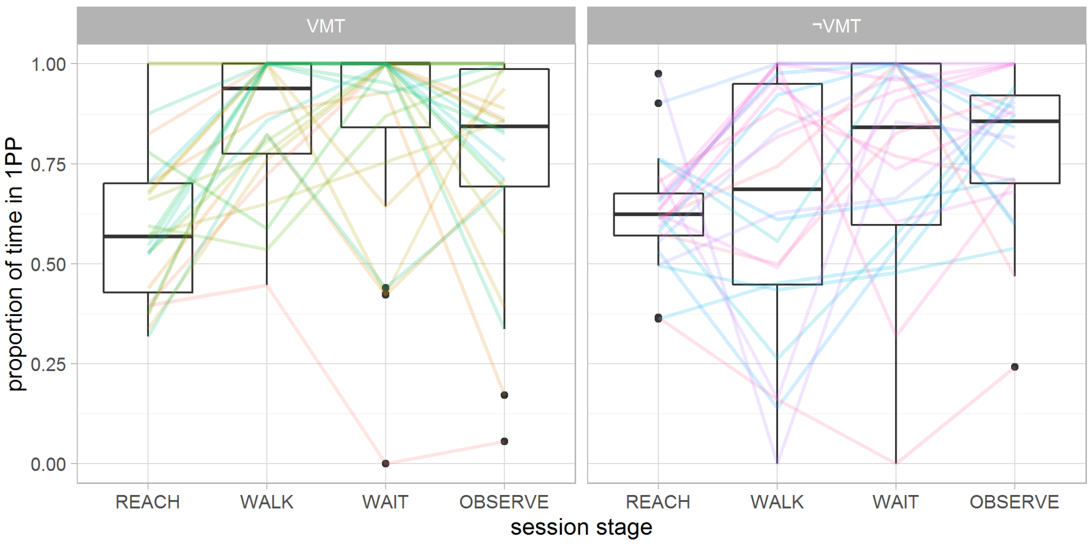


Figure 4.18 – Proportion of time spent in 1PP for each stage of the ALT session. Subjects tended to make a balanced use of perspectives in the REACH stage, while favoring 1PP for the following stages.

virtual body that 1PP seems able to promote. We also highlight that more subjects preferred the ALT condition, and that they had the perception of performing faster in that condition, even though we found no clear effect of perspective in our performance measure (Table 4.4). Moreover, none of the subjects reported feeling sick due to the perspective switch, although we highlight that no formal testing has been conducted to this sense.

Finally, the post experiment comparative questionnaire shows that subjects generally perceive the 3PP as safer than 1PP (Table 4.4). A potential application of ALT could be on post-traumatic stress disorder and phobia treatment, in which one can develop a strong sense of embodiment of the virtual body in 1PP, and eventually switch to 3PP when the body is exposed to a threat. This would allow the exposure to happen in a more reassuring manner, while still preserving a stronger bound to the virtual body, thus making the experience of self-exposure gradual.

4.4 Synthesis and Conclusion

In the first experiment we found a strong main effect of visuomotor synchrony to all questionnaire items, but no main effect of perspective. Reported senses of agency, ownership and self-location were high as long as the sensorimotor response was synchronous. Moreover, 3PP reduced search time of targets in the vicinity of the virtual body, while 1PP allowed more precision when approaching the target. These performance results were expected, given that, as the 3PP gives immediate feedback of user surroundings, but reduce angular size and stereoscopy view of targets. Interestingly, we also found that subjects act similarly in the reaching task, independently of perspective.

Table 4.4 – Post-experiment responses for the VMT group. Most subjects preferred to use 1PP, and felt safer in 3PP. When asked about conditions, subjects though ALT to be more efficient in the reaching task. ALT was also preferred by more subjects than the other conditions.

Which point of view ...	1PP	3PP	
... makes you feel safer when the floor falls?	3	**21**	
... do you prefer to use when the floor falls?	**19**	5	
... do you prefer to use to walk forward?	**22**	2	
... do you prefer to use to reach the targets?	**19**	5	

Which condition ...	1PP	3PP	ALT
... do you prefer to perform the reaching task?	2	2	**19**
... is more efficient to reach the targets?	8	5	**10**

In the second experiment, we saw a main effect of visuo-motor-tactile congruence on agency and self-location, but not on the sense of ownership. In turn, perspective had a main effect on self-location, threat and GSR, which were overall lower for 3PP. Interestingly, we also found an interaction between these factors for the agency and ownership responses. It yields similar agency and ownership scores across perspective conditions when multimodal congruency is present, but lower scores for 3PP when there is no multimodal congruency. We have also tested the ALT condition, which offers a new take on how switching perspective can feel natural while permitting the experience of a 3PP viewpoint on demand. Responses for the ALT condition were similar to 1PP regardless of the number of times that subjects switched perspectives or the proportion of time that they spent in each point of view.

Our studies diverge from literature on the sense of embodiment in 3PP due to the fact that we not only manipulate aspects of visuo-tactile congruence, but also allowed full-body control (visuo-motor congruence), including the global aspects of the body by walking. This is important because recent evidence on the relative contribution of visuo-motor and visuo-tactile congruencies suggests the predominance of the former, as well as an additive effect [Maselli and Slater, 2013, Kokkinara and Slater, 2014].

Although the sensorimotor manipulation was not equivalent across experiments, both experiment presented similar outcomes in the congruent condition. However, results for the incongruent conditions diverge. On the one hand, the sense of embodiment in 1PP was strong independent of the sensorimotor contingencies for the second experiment. On the other hand, sense of embodiment in 1PP was affected as much as for 3PP when a second of latency was added to the motion capture data in the first experiment. Notably, it suggests that 1 second of latency may be more detrimental to the sense of embodiment than no control of the body at all.

Some methodological differences between these two experiments have to be highlighted though. In the second experiment subjects were exposed to a virtual threat, and multisensory congruency was a between subject factor. In contrast, there was no threat in the first experiment, and the virtual environment was minimalist, with the visuo-motor synchrony as a within subject factor.

In summary, our results contribute to the understanding of the interplay of the multiple components supporting embodiment and show that several factors (visuomotor congruency, visuotactile congruency or perspective) can have a positive impact on body ownership and embodiment depending on the tasks to perform and on the stimuli provided. In our case, in absence (experiment 1) and presence (experiment 2) of tactile stimulation and in the context of action oriented tasks, visuomotor synchrony dominates over perspective. Under other circumstances, perspective can dominate over visuotactile congruency when the manipulation focuses on the contrast between the location of the touch and the change of perspective [Slater et al., 2010b]. Understanding the cognitive mechanisms of embodiment is a fundamental challenge for the development of VR interaction and needs to be investigated further. It is precisely because VR allows controlling factors such as perspective and analyzing behavior in ecologically valid conditions (e.g. differences in timing of reaching movements) that it provides the necessary environment for conducting this cognitive neuroscience research.

Table 4.5 – Perspective taking and embodiment statistical significance tests summary.

Variable	Multimodal Congruence	Perspective	Interaction
Agency	$F_{1,36} = 89.84$ **p < .001** $d = 2.47$	$F_{2,72} = 9.70$ **p < .001**	$F_{2,72} = 4.49$ **p < .015**
Ownership	$F_{1,36} = 4.49$ **p < .042** $d = 0.44$	$F_{2,72} = 22.75$ **p < .001**	$F_{2,72} = 5.22$ **p < .008**
Self-location	$F_{1,36} = 4.31$, **p < .046** $d = 0.41$	$F_{2,72} = 33.77$, **p < .001**	$F_{2,72} = 0.30$ $p > .738$
Threat	$F_{1,36} = 1.62$ $p > .075$	$F_{2,72} = 21.44$ **p < .001**	$F_{2,72} = 0.47$ $p > .627$
More bodies	$F_{1,36} = 3.84$ $p > .057$	$F_{2,72} = 4.34$ **p < .017**	$F_{2,72} = 6.76$ **p < .003**
Turning Virtual	$F_{1,36} = 0.00$ $p > .946$	$F_{2,72} = 16.41$ **p < .001**	$F_{2,72} = 0.74$ $p > .482$
GSR	$F_{1,36} = 0.59$ $p > .448$	$F_{2,76} = 4.21$, **p < .020**	$F_{2,72} = 0.28$ $p > .754$

Perspective post hoc

Variable	1PP vs. 3PP	1PP vs. ALT	3PP vs. ALT
Self-location	$t_{47} = 6.94$ **p < .001** $d = 1.52$	$t_{47} = 3.64$ **p < .001** $d = .57$	$t_{47} = 4.19$ **p < .001** $d = 0.54$
Threat	$t_{47} = 6.17$ **p < .001** $d = 1.04$	$t_{47} = 3.21$ **p < .003** $d = .48$	$t_{47} = 3.97$ **p < .001** $d = 0.50$
GSR	$t_{39} = 3.01$ **p < .020** $d = 0.37$	$t_{39} = 0.48$ $p > .630$ $d = .07$	$t_{39} = 2.47$ **p < .040** $d = 0.35$
Turning virtual	$t_{47} = 4.83$ **p < .001** $d = 0.87$	$t_{47} = 0.66$ $p > .510$ $d = 0.11$	$t_{47} = 4.53$ **p < .001** $d = 0.62$

Interaction post hoc

		VMT	
Variable	1PP vs. 3PP	1PP vs. ALT	3PP vs. ALT
Agency	$t_{23} = 1.16$ $p > .777$ $d = 0.32$	$t_{23} = 0.47$ $p > .948$ $d = 0.13$	$t_{23} = 0.73$ $p > .948$ $d = 0.19$
Ownership	$t_{23} = 2.60$ $p > .079$ $d = 0.69$	$t_{23} = 1.04$ $p > .928$ $d = 0.30$	$t_{23} = 1.25$ $p > .889$ $d = 0.28$
More bodies	$t_{23} = 4.56$ **p < .002** $d = 1.00$	$t_{23} = 2.55$ $p > .124$ $d = 0.63$	$t_{23} = 1.32$ $p > .601$ $d = 0.34$

		¬VMT	
Variable	1PP vs. 3PP	1PP vs. ALT	3PP vs. ALT
Agency	$t_{23} = 3.69$ **p < .008** $d = .94$	$t_{23} = 1.46$ $p > .627$ $d = 0.30$	$t_{23} = 3.52$ **p < .01** $d = .68$
Ownership	$t_{23} = 5.48$ **p < .001** $d = 2.15$	$t_{23} = 3.18$ **p < .025** $d = 0.71$	$t_{23} = 4.30$ **p < .003** $d = 0.86$
More bodies	$t_{23} = 0.87$ $p > .601$ $d = 0.26$	$t_{23} = 1.92$ $p > .268$ $d = 0.40$	$t_{23} = 2.16$ $p > .249$ $d = 0.53$

	1PP	3PP	ALT
Variable	VMT vs. ¬VMT	VMT vs. ¬VMT	VMT vs. ¬VMT
Agency	$t_{37} = 4.78$ **p < .001** $d = 1.97$	$t_{38} = 8.32$ **p < .001** $d = 3.26$	$t_{42} = 6.27$ **p < .001** $d = 2.17$
Ownership	$t_{45} = 0.22$ $p > .928$ $d = 0.06$	$t_{46} = 3.01$ **p < .024** $d = 0.93$	$t_{43} = 1.01$ $p > .928$ $d = 0.26$
More bodies	$t_{46} = 3.12$ **p < .026** $d = 0.88$	$t_{45} = 1.12$ $p > .601$ $d = 0.34$	$t_{45} = 2.06$ $p > .249$ $d = 0.56$

5 Movement Distortion and Embodiment

In this Chapter we use VR to quantify the extent to which a subject self attributes a distorted movement. Specifically, the movement of an arm that is manipulated in order to facilitate or hinder the completion of a reaching task. We achieve this by decreasing or increasing the amplitude of the hand movement required to reach for a target, while maintaining the apparent amplitude (visual feedback) fixed. Thus, the most salient feature of the distortion during the movement is that the visual feedback may move faster or slower than the real (performed) movement. This builds into a visuo-proprioceptive discrepancy, characterizing a spatiotemporal distortion.

We perform two experiments with our distortion model. The first aims at quantifying the limits of self-attribution of the distorted movement, in which subjects were asked if a seen movement matches the movement they have performed. The second experiment acquires subject's impressions on whether a given level of distortion makes the reaching task easier or harder to complete than expected. The latter is not obvious because it involves a trade-off between the manipulated movement amplitude (objective manipulation of difficulty) and the subjects' capacity to promptly correct an ongoing movement that has been distorted. This topic is further detailed in Section 5.2.

In our context of research, the spatiotemporal distortion can be used to facilitate or hinder the completion of a task, to prevent interpenetration with other elements of the VE [Burns et al., 2006, Burns et al., 2007], or to accommodate the visual surface of a virtual object into a passive haptics device of different shape [Kohli et al., 2012, Kohli et al., 2013]. For instance, according to user's engagement and their current level of ability, an application designed for physical activity can redirect the virtual body movements in order to reduce or augment the effort necessary to complete a task. This could be the case for applications on physical rehabilitation, such as for post stroke patients, who may experience reduced mobility and impaired fine control of movements [Rohrer et al., 2002].

Moreover, the interpretation of such manipulation includes questioning how the interplay of discrepancy detection and judgment of agency influences the self-attribution of a movement.

We discuss it in terms of current theories of agency.

5.1 Related work

While aspects of motor planning have been discussed in Section 2.3, here we present relate work on the context of VR.

Burns has explored two aspects of visuo-proprioceptive mismatch. The first on the perception of miss-location of real and virtual hand, [Burns et al., 2006] shows that a person may be strikingly unaware of visuo-proprioceptive mismatches which were gradually introduced over a long period of time. Specifically, subjects that have been primed to know the mismatch would happen only notice the discrepancy when it reaches ≈ 20 deg, while unprimed subjects would only realize it when the mismatch reaches ≈ 40 deg. The second aspect concerns the perception of movements with spatiotemporal distortions (speed reduction/amplification, results presented in Table 5.1) [Burns and Brooks, 2006]. They finally propose an interaction technique that takes advantage of both perceptual limits to prevent interpenetration of the virtual hand with the (immaterial) VE.

In a related topic, [Kohli, 2010] explores the distortion of movements in order to redirect haptic sensation. The goal is to use a passive haptic device as a proxy to a more complex virtual object. To evaluate this concept the authors designed an experiment where subjects had to perform a multi-directional pointing task over a tilted plane, while the visual feedback presented an user aligned tapping plane [Kohli et al., 2012]. The authors have aimed at identifying potential performance aspects, and have only informally evaluated the subject of perception [Kohli et al., 2013].

This specific topic has also been explored by another group, mostly relying on augmented reality setups. In [Ban et al., 2012a], the authors redirect haptic feedback of complex symmetric objects to a proxy cylinder. This work was further extended to represent more complex symmetrical objects by adding bumps to the proxy object [Ban et al., 2012b]. Finally, They explore the inclusion of more than one point of redirection (normally the index finger), to allow for pinching gestures [Ban et al., 2014].

The work on redirected haptics has been further explored by [Spillmann et al., 2013] on the field of surgery simulation. In this case, the remapping was relevant to understand the visuo-motor-tactile responses one receives through a surgical instrument (i.e. tool mediated contact preserving).

Furthermore, [Kokkinara et al., 2015] have shown that by increasing the speed of reaching movements, an after effect change to the perception of space could be observed. More specifically, after being exposed to a spatiotemporal distortion (2x and 4x the speed) between real and virtual hands (avatar), subjects tended to overestimate the size of an object, indicating a visuo-proprioceptive remapping. They have also shown that this distortion seems to have only a small impact to the sense of embodiment of the virtual body.

Finally, [Lecuyer et al., 2000, Lécuyer et al., 2008, Jauregui et al., 2014] explores the notion of pseudo haptics, which examines cross-modal perception in order to create the subjective sensation of haptic interaction with objects of different physical properties. For instance, [Lecuyer et al., 2000] manipulates the control-display ratio (CDR - the ratio mapping the input of a device to an output in a display) of a mouse to convey pseudo haptic sensations. The mouse is used to control a cube in the screen, and when the cube passes through a delimited area, the CDR could either increase or reduce. Subjects reported the sensation of "lightness" and "gliding" when the CDR increase, and "friction" and "viscosity" when the CDR was reduced. That is, the added/reduced effort resulting from the longer/shorter distance the subject had to cover due to an incongruent visual, proprioceptive and tactile feedback was felt as a tangible obstacle.

5.2 Distortion Model

We implement a distortion model that alters the visual feedback of movements in order to facilitate or hinder a reaching action. In practice, we reduce or increase the distance between the position where the movement starts and the target.

Figure 5.1 graphically depicts the behavior of the function in 1D, with the target position ($\mathbf{p}_{tgt}$) equals to 0, we assume that the target is static. The horizontal axis represents the real hand position ($\mathbf{p}_{real}$), while the vertical axis depicts the redirected (virtual) hand position ($\mathbf{p}_{virt}$). The movement is mapped with a 1:1 ratio while it is outside a given distance range ($d_{range} = 1$ around the target in the figure). Once it enters the d_{range}, the facilitating distortion (green lines) speeds up the movement until the virtual hand reaches the target (position 0), while the hindering distortion (red lines) slows down the movement. Conversely, once the virtual hand reaches the target the movement speed is inverted. As a result, the interaction size of the virtual target is not altered and the virtual hand is brought back to collocation with the real hand. The black line represents the movement without remapping.

The position of the virtual hand is defined by:

$$\mathbf{p}_{virt_n} = \begin{cases} \mathbf{p}_{real_n} + \hat{\mathbf{v}}_{dir_n} \times \Delta d_n, & if\ ||\mathbf{p}_{tgt} - \mathbf{p}_{real_n}|| < d_{range} \\ \mathbf{p}_{real_n}, & otherwise \end{cases} \tag{5.1}$$

where $\mathbf{v}_{dir_n}$ is the direction of the distortion

$$\mathbf{v}_{dir_n} = \begin{cases} \mathbf{p}_{tgt} - \mathbf{p}_{real_n}, & if\ ||\mathbf{p}_{tgt} - \mathbf{p}_{real_n}|| > d_{range} \\ \mathbf{v}_{dir_{n-1}}, & otherwise \end{cases} \tag{5.2}$$

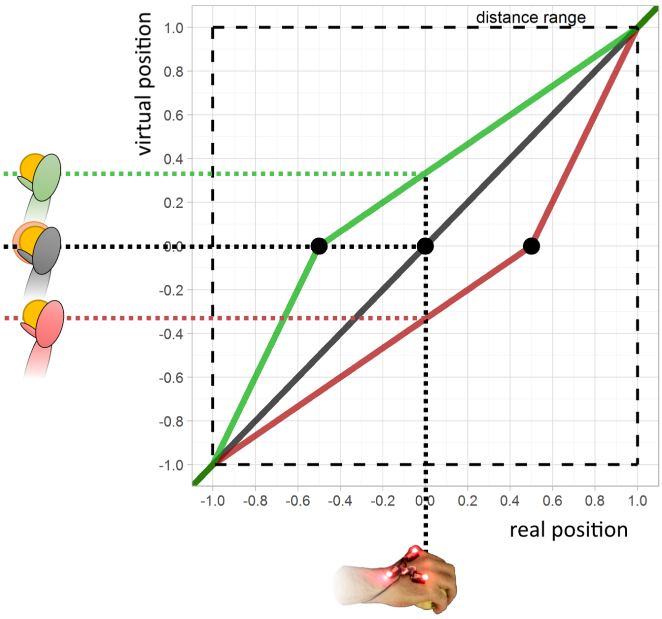


Figure 5.1 – Overview of the reaching distortion function. Horizontal axis depicts the real hand position, while the vertical axis depicts the virtual hand position. The lines map a movement from the left to the right, with a distance range of 1 and a target position set at 0. The virtual colored hands in the left shows where the current real hand position would be mapped into. The green and red colors represents a facilitating and a hindering distortion respectively.

and Δd_n describes the magnitude of the distortion

$$\Delta d_n = \left(\frac{||\mathbf{p}_{tgt} - \mathbf{p}_{virt_{t_{n-1}}}||}{d_{range}} - 1\right) \times d_{gain} \times d_{range} \tag{5.3}$$

Finally, d_{gain} defines the difference in distance proportion (normalized according to d_{range}) that the subject has to cover in order to reach for the target. If $d_{gain} = 0.5$, the movement becomes 50% longer than the apparent distance. If $d_{gain} = -0.5$, the movement becomes 50% shorter than the apparent distance. Based on the definition of index of difficult (ID), we expect this distortion to alter the difficulty of the reaching task. The ID is normally used as part of the Fitts law (a reaching/pointing time prediction equation often used in the field of human computer interaction). ID is defined by the equation:

$$ID = \log_2\left(\frac{D}{W} + 1\right) \tag{5.4}$$

Where D represents movement amplitude (distance) and W the target width. Intuitively, increasing the distance to the target (D) results in a higher ID, while reducing this distance results in a lower ID, modulating the ID of the task while the W parameter is kept constant. Experiment 2 evaluates whether this distortion effectively alters the difficulty of the task (Section 5.5).

The most salient feature of our distortion model *during* the movement is the difference in velocity. A distortion that facilitates the reaching movement presents an increased velocity until the virtual hand reaches the target, and a reduced velocity if the movement continues on the same direction, until virtual and real hand matches in position by leaving the distance range. The opposite happens with a distortion that hinder the movement. In the experiments we perform, the first part of this movement is always present. Thus, we opted to set the distortion in terms of change in speed, instead of the change in movement amplitude. This way we are in line with the hypothesis that it is a mismatch between the actual and predicted (by forward models) sensorial input that brings an error to awareness of the subject [Blakemore et al., 2002]. To do so the variable d_{gain} can be defined by a *speed multiplier* using $d_{gain} = -\frac{speed_{mult}}{speed_{mult}+1}$.

5.3 Materials and Methods

Equipment and Software

An Oculus development kit 2 HMD was used to display the virtual scene (960 x 1080 pixels per eye, 100 deg field of view, 75 Hz). Head tracking was performed using its inertial sensors (low latency) and corrected for drift around the vertical axis using optical tracking.

A pair of Bose® Quietcomfort 15 headphones were used for environmental noise canceling and to provide non localized white noise, thus phonically isolating the user from the real environment.

A PhaseSpace ImpulseX2 optical tracking system was used for motion capture. Our setup uses 18 cameras and a total of 14 LED markers, from which 4 were attached to the HMD and 10 to the upper limbs of the subject. Three markers were fixated in a non-collinear arrangement over the back of each hand of the subject, allowing the reconstruction of position and orientation of the hand in free space (6 degrees of freedom). A marker was fixated over the top of each shoulder, allowing to track trunk movement to some extent. Finally, a marker was fixated to each elbow, these markers are used to solve the ambiguity of elbow bend direction relative to the shoulder to hand vector. Figure 5.2a shows a subject wearing HMD, headphones and LED markers. We assessed a latency in the range of $30ms$ to $40ms$ from physical action to HMD display.

The virtual environment was developed using the Unity game engine. It consists of a virtual body, a chair and a carpet – collocated with the subject body and a real chair and carpet used in the real environment.

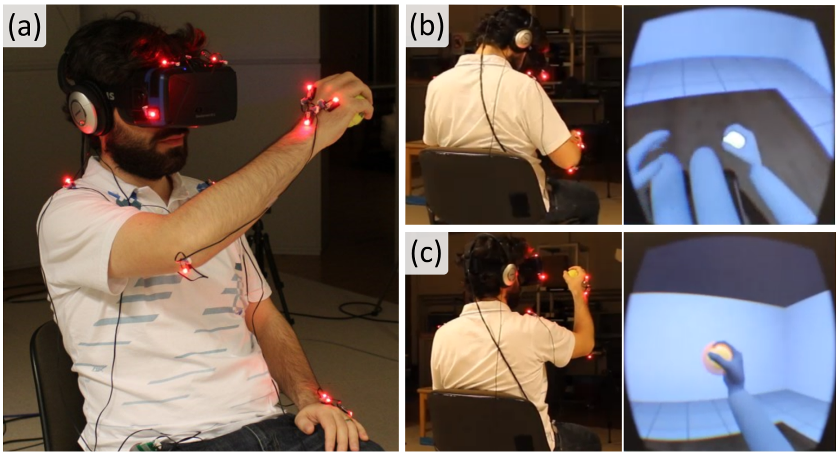


Figure 5.2 – Overview of the reaching experiment setup. (a) shows a subject equipped with a HMD, headphones and motion capture markers. In (b) the subject observes the virtual body, and in (c) he performs the task.

The virtual body was animated in real time using the FinalIK[1] package available at the Unity asset store[2]. Virtual body hands position and orientation have high priority, and the rest of the posture is defined ensuring the collocation of the virtual hands with the rigid body defined by the LED markers (plus the position deviation). Hips and legs were not tracked nor animated, participants were asked to remain seated during the whole VR blocks of the experiments. Figure 5.2bc shows sample captures of the posture reconstruction used in the experiments. Moreover, the virtual body and its limbs were scaled to approximately match the body of the subject. This was done before the start by measuring the subjects height and their right arm and leg segments.

As the markers attached to the hands could not be fixed in identical positions across subjects, a short real hand to virtual hand registration was necessary. Once equipped with the HMD, the participant could see small green spheres at positions corresponding to the LED markers being worn. The participant was then asked to position these spheres over the hands of the virtual body, at the equivalent position where the LED markers are located in her hand.

Task

Participants had to repeatedly perform a reaching task during both experiments. The task consisted of two movements and a question. In the first movement the subject had to take a tennis ball hold with their right hand inside a semitransparent virtual target. After a random interval lasting between $200ms$ and $600ms$ inside that target, the target disappeared and a second semitransparent target appeared in a position opposite to the first. The participant had to perform a second movement and take the ball inside the new target. However, the visual

[1] root-motion.com
[2] assetstore.unity3d.com

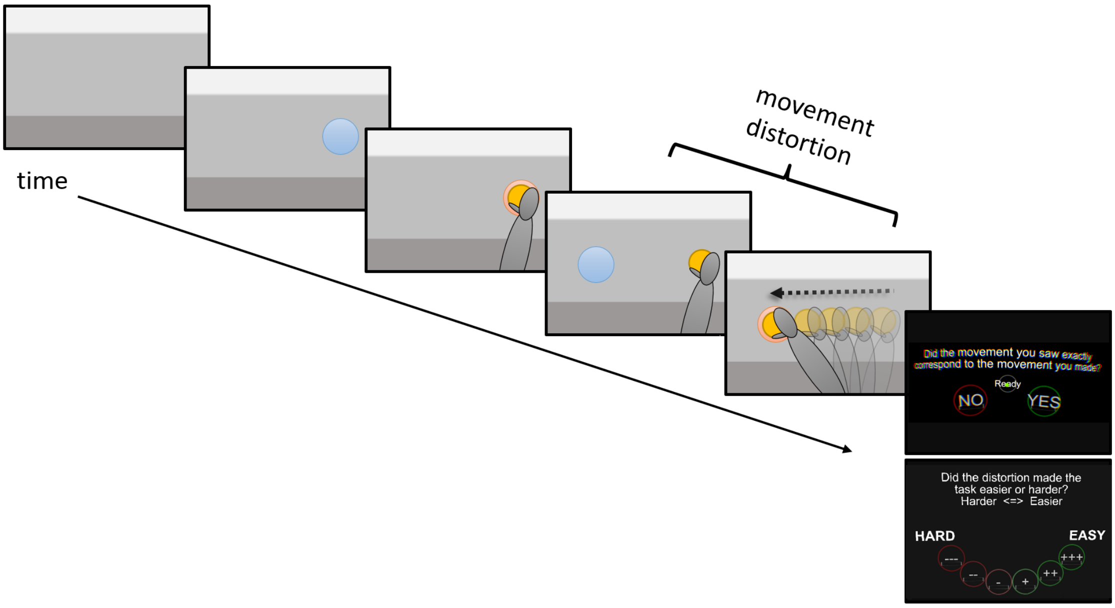


Figure 5.3 – Overview of the reaching experiment trial. The trial consisted of two movements and a question: the first movement uses a target to position the hand and is not distorted; the second movement goes from target one to target two and may or may not be distorted. The question was different for each experiment: experiment 1 asked whether the movement was exactly like the one performed by the subject; experiment 2 asked if the applied distortion made the task easier or harder than with no distortion.

feedback of the second movement could be distorted (spatiotemporal distortion), interfering with the task. The reaching is complete once the ball is kept inside this second target for $150ms$. The tennis ball and the semitransparent targets have a diameter of $\approx 6.7cm$ and $10cm$ respectively. Finally, a forced choice question appears, participants could answer the question by orientating their head to face the desired answer. The question and answer options were different according to the experiment. Figure 5.3 gives an overview of a task trial. The subject had to lower the right hand before the next trial could start.

We used 4 predefined positions for the targets. They were arranged around a central point at the height of the eyes of the subject. One target above, one below, one to the left, and one to the right of this point. These four positions defined a plane parallel to the projection plane of the virtual cameras (assuming that the subject faces forward). The distance of each target from the central point was equivalent to 25% of the subject's arm length.The central point was at 50% the subject's arm length far from the camera. Therefore, if a the task defines a trial with a distorted movement leftwards, the subject had to first reach for the target to the right of the central point, and then perform a movement to the target located to the left of the central point. The apparent distance of the movement is equivalent to 50% of the subject's arm length, while the actual movement also depends on the magnitude of the distortion defined by d_{gain}. The positions of the central target and targets are defined relative to the head position of the subject at the start of each block.

Procedure

The subject was first asked to read an information sheet and to complete and sign an informed consent form. Next the subject was asked to fill in a characterization form, with regular questions about their background (other experiments, experience with HMDs ...) and physical characteristics (height, weight and age).

Then, the experimenter measured the approximate length of the right arm, forearm, thigh and leg of the subject, which was used to scale the virtual body dimensions to match those of the subject. This information was also used to set the distance between targets, so that the effort and the ability to reach could be kept roughly equivalent across subjects. The subject is instructed to sit and the experimenter then fixates the optical markers to both arms of the subject. Once ready, the experimenter gives a detailed overview of the stages of each trial, and expose the structure of the experiment, which starts with two short training blocks. Finally, the subject was equipped with the HMD and noise canceling headphones to start the training.

In the first training block the subject completed 8 trials without any redirection, two in each direction. In the second training block the subject completed 8 addition trials, but now with significant redirection (-.8 and .8), one for each combination of direction and distortion type. The subject was told beforehand whether there would be a distortion in the training block, and what answer the subject should give in such case. This procedure was adopted to ensure that the subject understood the task, and were shown what a movement distortion looked like without a verbal description of its features.

After completing the training the subject went through two blocks of trials, as described in Section 5.4. Each block took between 15 and 25 minutes, depending upon subject's pace and precision in recognizing a deviated movement. An interval was given between the blocks, as well as if the subject requested for a pause during the block. Experiment 1 was complete after the second block.

The subject was given time to rest before starting experiment 2 (Section 5.5). Once the subject was ready, the experimenter went through the new instructions. Experiment 2 consisted of 2 short blocks of trials (3 to 5 minutes), with movements always towards the left of the subject. Finally, the experimenter conducted a debriefing with the subject.

A total of 20 subjects participated on both experiments (mean age 23.9 with SD of 4.5, 3 female). Six subjects reported having participated in an experiment using virtual reality in the past, while 8 reported having tried a HMD in the past, one of which with weekly frequency.

This experiment was approved by the Commission cantonale d'éthique de la recherche sur l'être humain in Vaud, Switzerland. Subjects signed a consent form and were paid 20 CHF/hour for their participation.

5.4 Experiment 1: Just Noticeable Difference

Experiment 1 was designed to estimate the limits of subjective self-attribution of a redirected movement that facilitates or hinder the completion of a goal directed task. After the completion of each trial, we ask the subject whether "the movement you saw exactly corresponds to the movement you have performed", which the subject had to answer by facing a "Yes" or "No" timed button. The limits of self-attribution can be defined as the magnitude of distortion to which the subject as likely to self-attribute a distorted movement (answer "Yes") than not.

The experiment followed a within subject design with two independent variables: orientation of distortion (facilitating or hindering movement) and movement direction (left, right, up or down).

To quantify these limits we adopt concepts and procedures from psychophysics. Psychophysics acts on the understanding of how a stimuli affects one's sensation/perceptions, its methods are often employed to assess the just noticeable difference (JND) between a standard and an altered stimuli. The JND can be interpreted as a constant proportion K of the intensity I of the standard stimulus, as defined by Weber's law:

$$\Delta I = K * I \tag{5.5}$$

Thus, we focus on measuring the constant K, which is then used to compute the ΔI for a given stimulus intensity I. In our case, once we know what is an admissible K for our distortion (one for facilitating and one for hindering) we want to compute a *virtual* stimulus given a *real* stimulus:

$$I_{virtual} = I_{real} + \Delta I = I_{real} + I_{real} * k = (1 + k) * I_{real} \tag{5.6}$$

Moreover, our study is distinct from regular psychophysics paradigms in two ways:

(i) We assess the JND across different sensory modalities, i.e. the visual feedback is altered with regard to its forward model prediction and proprioceptive feedback;

(ii) The question is not explicit about the physical features of the distortion, instead it asks whether the subject consider the movement they see to be equivalent to the one they have performed.

The speed change is the most salient stimuli, and its proportion constant is defined by the speed multiplier $speed_{mult}$. However, based on preliminary tests and related work [Burns and Brooks, 2006], we observed that the $speed_{mult}$ yields strong asymmetry between the mean

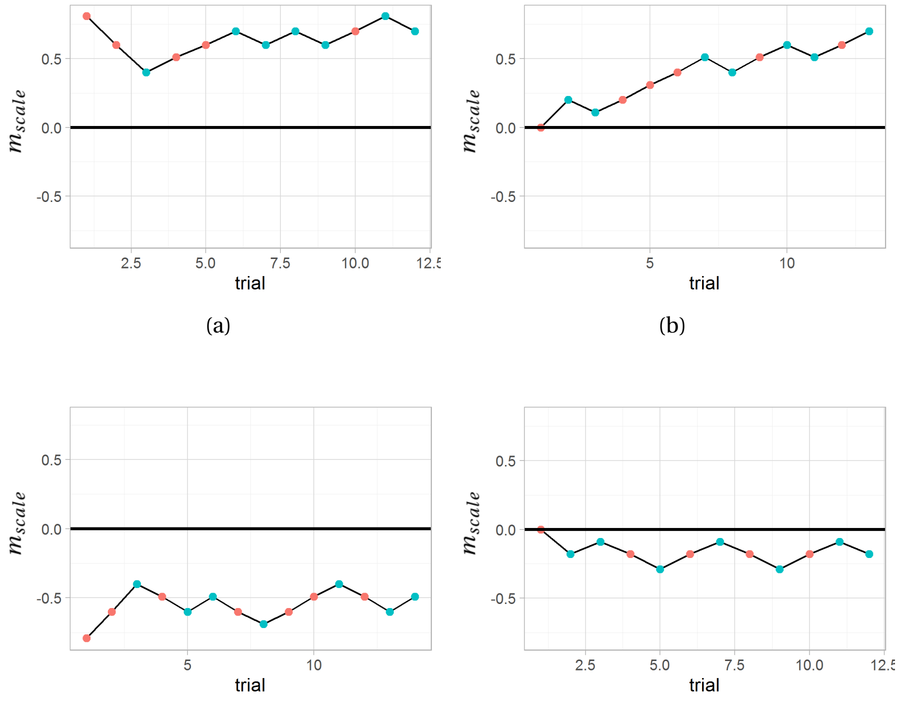


Figure 5.4 – Samples of adaptive staircases. Note that the staircases were oriented to either reduce (ab) or increase (cd) the required movement, starting either from a high distortion (ac) or a no distortion (bd) value. The JND was computed as the mean of the last 4 blue points, which represents turns in the trend of the staircase.

thresholds and variance of facilitating and hindering distortions. Thus, we decided to define the measurement intervals of the experiment in terms of $m_{scale} = \log_2(speed_{mult} + 1)$. As such, the measurement scale can be converted back into speed with $speed_{mult} = 2^{m_{scale}} - 1$, for which the m_{scale} values of -1, 0 and 1 corresponds to $speed_{mult}$ values of -0.5, 0 and 1 respectively. That is, -1 is half the speed, while 1 is twice the speed.

In order to assess the JND, distortion intensity was controlled with an adaptive staircase, a procedure that changes the intensity of the stimulus discrepancy based on the whether the subject identified or not the discrepancy in the last trial [Meese, 1995]. In our specific case, if the participant answers "Yes" to a correct or distorted movement, the discrepancy is increased. If the subject answers "No" to a distorted movement, the discrepancy is decreased. Finally, if the subject answers "No" to a correct stimuli, we do not alter this parameter, as this would change the orientation of the distortion (from facilitating to hindering and *vice versa*).

The staircase was complete when either the subject changed the direction of the staircase 7 times (e.g. from a distortion increase to a distortion decrease trend) or performed a total of 20 trials in the same staircase. The JND was computed as the mean of the 4 last staircase turns (Figure 5.4). Each subject underwent a total of two blocks of 16 staircases, two for each combination of movement direction and distortion type, for a total of 32 staircase procedures. Thus, for each combination of distortion orientation and movement direction the subject performed 4 staircases: 2 starting with correct movements, with an initial trend of distortion increase, and 2 starting at a high level of distortion, with a distortion decrease trend (Figure 5.4).

The size of the staircase step changed dynamically, it starts as 0.2 (-0.2 if the trend is to decrease the distortion), and after the first staircase turn it is reduced by half. This value is kept for the rest of the trials of that staircase. To prevent the sequential presentation of trials from the same staircase, 4 of the 16 staircases of the block are run concurrently.

A second relevant measurement is the point of subjective equality (PSE), i.e. the point at which the participant subjectively evaluates a presented stimulus to be equivalent to the standard one. The PSE may be computed as the point in between the facilitating and hindering JNDs.

A sequence of 8 non distorted movements was presented in the start and end of each block, for which the subjects were not aware. We expected to observe if an adaptation could occur by assessing the change in the rate of "Yes" answers to the non-distorted movements preceding and succeeding the experimental block.

5.5 Experiment 2: Task Difficulty

The second experiment acquires subject's impressions on whether a given level of distortion makes the reaching task easier or harder to complete than it would be without any distortion. For instance, we suppose that the difficulty will change according to the index of difficulty (ID), as used by the Fitts law. Intuitively, the ID increases if a distortion imposes a bigger amplitude of movement (hindering distortion), and decreases if a smaller amplitude is imposed (facilitating distortion). However, the distortion might also cause a big mismatch between internal forward models predictions of sensorial input and actual sensorial input, requiring the subject to promptly adapt an ongoing movement in order to comply with the distortion. Moreover, behavioral experiments have shown that the minimum delay needed for a visual or proprioceptive signal to influence an ongoing movement is 80–100 ms [Desmurget and Grafton, 2000]. Thus, if a movement is shortened by too much it may become unpractical in terms movement control mechanisms, potentially contradicting the assumption we make with the index of difficulty.

Experiment 2 followed a factorial within-subject design, with distortion intensity as the only independent variable (in the same scale as experiment 1, with 9 values ranging from $-.8$ to $.8$ in steps of $.2$). Movement direction was always towards the left.

The response variable was the difference in difficulty, after each reaching the subject was asked: "Did the distortion made the task easier or harder?". The answer was given in a 6 points

Table 5.1 – Estimated points of Just Noticeable Difference (JND) for different scales (Mean ± Standard Deviation), and comparison with experiment from literature.

Direction	JND [$speed_{mult}$]* [Burns and Brooks, 2006]	JND [$speed_{mult}$]	JND [m_{scale}]	t-test
	faster | slower	faster | slower	faster | slower	$p <$
Left	$+.44 \mid -.08$	$+.86(\pm.38) \mid -.13(\pm.07)$	$+.82(\pm.27) \mid -.21(\pm.12)$	**.001**
Right	$+.40 \mid -.06$	$+.84(\pm.39) \mid -.21(\pm.06)$	$+.83(\pm.29) \mid -.36(\pm.12)$	**.001**
Up	$+.51 \mid -.16$	$+.65(\pm.34) \mid -.18(\pm.06)$	$+.68(\pm.29) \mid -.29(\pm.10)$	**.001**
Down	$+.38 \mid -.27$	$+.90(\pm.44) \mid -.27(\pm.06)$	$+.85(\pm.29) \mid -.47(\pm.11)$	**.001**

*Values from [Burns et al., 2006], in which the task was not target directed and the question explicitly concerned speed perception.

scale (Fig 5.3). Participants were led to believe that all the trials were distorted.

The experiment was divided into two short blocks, each with a total of 36 trials, 4 for each of the 9 levels of distortion intensity. By asking a more explicit question about the distortion, we expect to find whether the subject is capable of perceiving that a distortion is presented, and to consistently rate its difficulty (e.g. not using the center of the difference in difficulty rating scale).

5.6 Results

For the analysis of experiment 1, we exclude the staircases that failed to converge, we defined these as: (i) staircases that reached 20 trials before completing a minimum of 5 turns, (ii) answering "no" when the staircase was at a no distortion point in the range of the last 4 staircase turns. A total of 640 staircases were completed, 62 of which removed because of (i) and 39 because of (ii), leaving a total of 539 staircase procedures. We also excluded 2 subjects from further analysis, as they failed to converge in at least 1 staircase per combination of conditions.

The JND results in terms of $speed_{mult}$ and m_{scale} are presented in Table 5.1, absolute values comparing the distortion orientation is presented in Figure 5.5. We also present the JND in terms of percent of reduction (facilitate) and increase (hinder) of the required movement amplitude (i.e. d_{gain}) in Table 5.2 and Figure 5.5. When using d_{gain} for the scale we obtain similar variances, i.e. a balanced distribution across the facilitation and hindering distortions. Therefore we decided to compute PSE in this scale (the mean of both JND).

For experiment 2, results are presented in Figure 5.6. The blue line and the shaded region represents a loess (locally weighted regression) fit and its 95% confidence interval [Cleveland and Devlin, 1988]. The vertical dashed lines represents the results of experiment 1 for the leftward movement for comparison. The green and red shaded areas highlight the "easier" and "harder" levels of difficulty available in the scale.

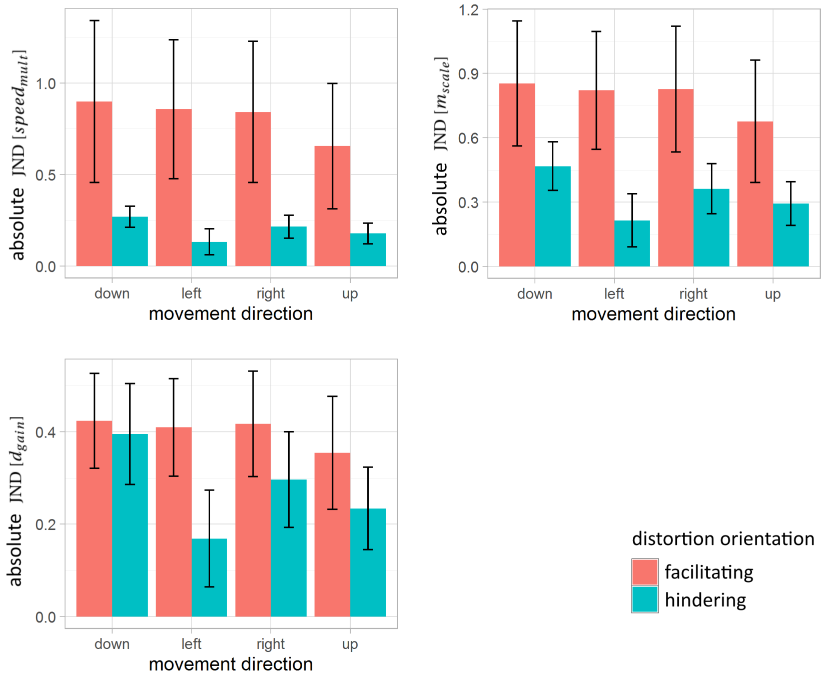


Figure 5.5 – Bar plots of the JND thresholds in different scales. The difference between facilitating and hindering movement was significant for all directions and in the different scales, except for the downward movement in the change in amplitude scale.

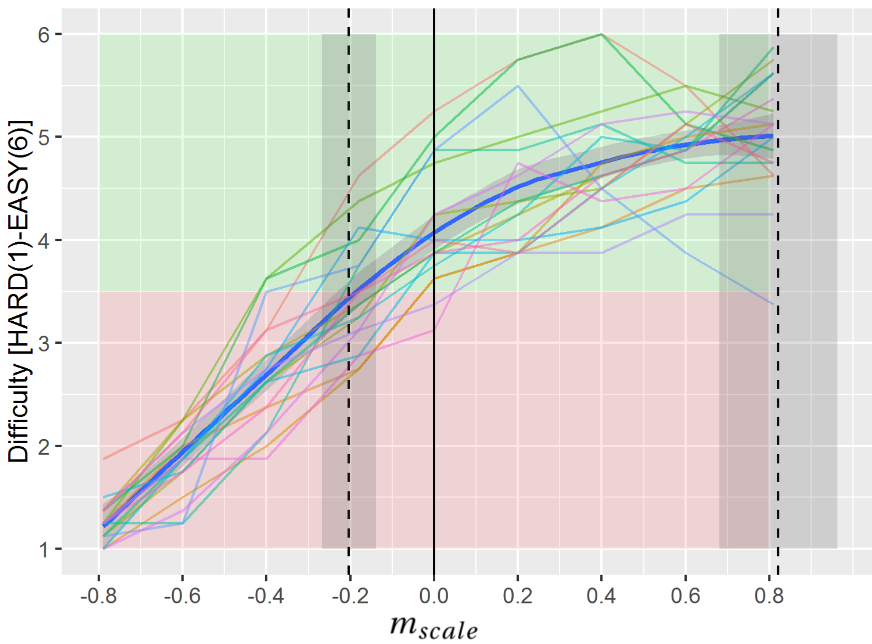


Figure 5.6 – Subjective evaluation of the difference in difficulty due to movement distortion. The point where subjects become uncertain of whether the distortion was affecting difficulty coincides with the JND for hindering distortion. Note that in this experiment the subject only performed movements toward the left.

Table 5.2 – Estimated points of Just Noticeable Difference (JND) and Point of Subjective Equality (PSE) for distance proportion difference scale (d_{gain}, Mean ± Standard Deviation).

Direction	JND [d_{gain}] decrease | increase	t-test $p <$	PSE [d_{gain}] arithmetic mean	PSE [d_{gain}] gaussian fit mean
Left	$-.41 \pm .11$ | $+.17 \pm .10$	**.001**	$-.120 \pm .07$	$-.116$
Right	$-.42 \pm .11$ | $+.30 \pm .10$	**.004**	$-.060 \pm .08$	$-.070$
Up	$-.35 \pm .12$ | $+.24 \pm .09$	**.008**	$-.060 \pm .08$	$-.077$
Down	$-.42 \pm .10$ | $+.39 \pm .11$	.343	$-.014 \pm .06$	$-.013$

5.7 Discussion

The JNDs we have obtained are higher than the closely related work of [Burns and Brooks, 2006]. This was the case especially for the facilitating distortion, which was more than 2 times bigger for two of the 4 movement directions. The higher tolerance might reflect two factors: we do not prime the subject to look for a specific physical feature of the distortion; and we ask about the experience of agency (self-attribution). This difference of the JND in the context of agency is valuable in the research of embodied interaction. In such case we are not focused on the limits of perception, but in the overall feeling of control of a body.

Moreover, we also emphasize that our task was goal directed, thus requiring a great level of attention and precision from the subject. Normally, one would expect discrepancies to be easier to spot under this condition, which was not the case if compared to [Burns and Brooks, 2006]. However, what became apparent is that subjects are biased to self-attribute movements *as long as* the task becomes easier than its apparent difficulty.

Although we might not have direct access to the output of predictions made by the forward models, the comparator mechanism suggests that if sensorial input and predictions mismatches are big enough, one may become aware of the sensory discrepancy [Frith et al., 2000, Blakemore et al., 2002]. However, below a certain threshold, the brain will typically monitor the movement, and may correct for visuo-proprioceptive discrepancies without the subject awareness [Nielsen, 1963, Jeannerod, 2003].

Curiously, when questioned about the means used to identify if a distortion did occurred, subjects often reported using the effort, e.g. expected and dispensed effort to complete the reaching task. This contradicts the use of an online comparator model, suggesting that the self-recognition in an action (or lack thereof) was often the result of a retrospective component of agency [Haggard and Chambon, 2012]. This approximates our results to the account of agency proposed by [Synofzik et al., 2008b], in which a higher order – non-minimal – representation of the self and its current state and intentions can affect how one evaluate the ownership of actions.

Moreover, the perception of hindering distortions across movement directions is not as uni-

Table 5.3 – Potential bias according to distortion orientation and movement direction.

Direction	Distortion type	potential bias
Left	Facilitating	-
	Hindering	arm reach limits
Right	Facilitating	-
	Hindering	-
Up	Facilitating	against gravity
	Hindering	against gravity
Down	Facilitating	toward gravity
	Hindering	toward gravity

form as the facilitating distortion. Notably, the movement toward left was especially sensible to the hindering distortion. We believe this relates to the different bodily receptors stimulated by the movements, e.g. the movement toward left requires the full extension of the arm, and involves self-contact with the chest. Moreover, the upward and downward movements are influenced by gravity. This might make the movement toward the right the less biased in terms of JND comparison (Table 5.3). Additionally, it is necessary to note that the measurement scale of the experiments (m_{scale}) are in $\log_2(speed_{mult}+1)$. Thus, the conversion to d_{gain} and $speed_{mult}$ may result in additional bias.

Nonetheless, experiment 2 validates the notion that our distortion model manipulates the difficulty of the task. It also suggests that subjects are capable of perceiving the facilitating distortion below the assessed JND interval when explicitly questioned about it.

Finally, the relation of distortion perception and effort that we have found suggests a link between our work and pseudo-haptics. Our manipulation is essentially similar to that of pseudo-haptics, it reduces/increases the control-display ratio of the arm movement. However, distinct from a computer mouse, the arm has an absolute relation with the body, and thus different sensory receptors. We believe that an interesting venue to investigate this relation is to understand how the thresholds we have found may apply to pseudo-haptics, e.g. if I want to produce the sensation of friction of the medium while moving the arm, are our thresholds capable of defining the minimal necessary distortion one has to apply?

5.8 Conclusions

In this Chapter we have explored the limits of self-attribution of a distorted movement in VR. Our distortion model allows continuity of movement, so that an end effector position is deviated and then attracted back into collocation with the real hand. We propose this behavior in order to preserve the width of the target, and also because we intend to use this distortion model in a more complex scenario in the future, such as the reaching scenario presented in Section 4.2.

We found that subjects can accept a wide range of distortions when questioned about self-identification of a movement. Notably, we found consistent evidence of a bias toward accepting distortions that make a task easier. Finally, the experiment on task difficulty suggests that, when asked directly about the distortion characteristics, subjects are aware about the facilitating distortion well below the JND we obtained with experiment 1. On the other hand, subjects could identify the hindering distortion as such in both experiments.

We believe that the thresholds we map here could be used in designing more engaging VR interactions. Particularly, movement distortion can be used to manipulate the difficulty of tasks, and consequently leveraging the challenge so that it matches the skills of the user and promotes the state of flow [Csikszentmihalyi and Csikszentmihalyi, 1992, Brondi et al., 2015]. Indeed, we envision its use in applications such as post stroke rehabilitation, in which the movement can be redirected to modulate the difficulty of the task, helping the subject to achieve the intended goal and gradually increasing the difficulty of the task as the subject progresses. Alternatively, by requiring more effort from healthy subjects, one could propose applications that instigate physical activity.

6 Conclusion

In this thesis we studied the sense of embodiment when bodily discrepancies are presented. Our main objective was to explore alternatives to prevent incongruent tactile feedback in virtual reality, while preserving a consistent sense of embodiment of a virtual body. That is, instead of extending the research in the direction of haptic feedback, we take a different perspective to the problem and explore two approaches that may help to prevent virtual body/environment contacts from happening:

- To provide a more informative take of the virtual body posture and its relation with the environment.
- To manipulate the visual feedback of movements in order to assess human sensitivity to postural distortion. These manipulations could be used to prevent visual artifacts (e.g. interpenetration).

In particular, we are interested on the impact of these manipulations on the senses of agency and ownership of the surrogate virtual body. For that we have drawn from the fundamental research on the embodied self that human perception is flawed, and that conflicting multi-sensorial stimulation may converge into an altered experience of bodily self-consciousness [Blanke et al., 2015].

To address the first item, we evaluated the use of non-planar projection in 1PP as compared to the use of 3PP (Chapters 3 and 4 respectively). We opted for the latter, preferring the specific setup in which the subject can alternate between 1PP and 3PP during the simulation.

To address the second item, we investigated the extent to which subjects accept/notice a distortion that remaps a performed movement into an incongruent visual feedback (Chapter 5). We find these limits to be rather high if compared to more objective perception measurements, especially when the distortion can facilitate the completion of the task.

We emphasize that we only look at a small set of manipulation possibilities within these two approaches. We further discuss this in the limitations and outlook sections of this Chapter.

6.1 Contributions

In this thesis we investigate alternatives to address two common limitations of embodied virtual reality, namely, field-of-view and the lack of physical feedback. We present three main contributions:

First, we manipulated visual feedback and arrived to a compelling option of alternating the point of view between 1PP and 3PP during simulation. Subjective evaluation of embodiment for this condition were very similar to those of 1PP alone, suggesting that the interruption of the point of view during the simulation is not detrimental to the experience of ownership of a virtual body. None of the subjects participating in the experiment reported sickness with the perspective transition that we propose, although we should point that no formal evaluation has been performed.

Second, we explore the influence of perspective taking and full body motor control on the embodiment of a virtual body. We show that the visuo-motor correlation over the whole body movement plays a strong role on the sense of ownership of a virtual body located in the extra-personal space (3PP). In the context of the scientific debates investigating the influence of perspective taking and visuo-motor contingencies over the sense of ownership, our result stands out by supporting the view that a 3PP is compatible with ownership. Therefore, we raise the discussion about the role of full-body task involvement (engaged interaction) in the sense of embodiment, which is not present in related literature, and could be related to the divergence of results. Moreover, although 3PP is competent at informing the relation between virtual body and its surroundings, our experimental results suggest a trade-off in terms of interaction reaching performance. This makes sense, while 3PP amplify the visible volume of the space, the angular size of end effector and targets are reduced, and stereo-graphic depth perception cues are reduced proportionally to the distance between observer and observed objects.

Third, we quantify the extent to which subjects tend to self attribute distorted movements (sense of agency). We have focused on the subgroup of goal oriented movements, more specifically, we interfere by adjusting the effort required to complete a task, consequently facilitating or hindering a goal directed movement. Our results show that subjects perform poorly in detecting discrepancies when the nature of the distortion is not made explicit. Additionally, we have found that subjects are biased toward self-attributing distorted movements that make the task easier. This is in line with two accounts of agency, the comparator model [Blakemore et al., 2002] and the notion of judgement of agency [Synofzik et al., 2008b].

Taking these contributions altogether, we can devise the following interaction design guidelines for embodied virtual reality:

1. Without long term adaptation, non-planar projections only hold an advantage for visual search time compared to regular perspective projection.

2. Non-planar projections could be further explored in an alternating interface, being used for quick inspection of the VE.
3. There is a performance trade-off between 1PP and 3PP for reaching. 3PP increases surrounding awareness but decrease visual size and occlude some targets.
4. Subjects make similar use of end effectors for selection regardless of the point of view of the virtual body.
5. Under optimal multisensory congruence condition – including sensorimotor contingencies – one might feel body ownership of a virtual body seen from a 3PP.
6. Alternating perspectives (1PP-3PP) did not affected the sense of embodiment as compared to 1PP, regardless of the multisensory congruence setting.
7. Considering the speed distortion scale, our self-attribution thresholds are considerably higher than the discrepancy perception thresholds reported in [Burns and Brooks, 2006].
8. The speed related measurement scale used in the movement distortion experiments may not be adequate considering the distribution of JND results. The change in distance is a more convenient scale, but the conversion might have added bias.
9. People seem to have a bias to self-attribute movements that make their task easier.

6.2 Outlook

Although it is the problem of the immateriality of the virtual world that bounds both of our courses of action together (perspective taking and movement distortion), the relevance of our research exceeds this boundary.

Particularly, movement distortion can be used to manipulate the difficulty of tasks, and consequently modulate the engagement of the user. Indeed, we envision its use in applications such as post stroke rehabilitation, in which the movement can be redirected to modulate the difficulty of the task, helping the subject to achieve the intended goal, and gradually increasing the difficulty of the task as the subject improves. Alternatively, by requiring more effort from healthy subjects, one may propose applications that stimulate physical movement.

Moreover, we believe that the investigation about tolerance to movement distortion can yield new venues to the development of motion capture hardware and animation software. The perception that the visual feedback can be severely altered without major effects to the self-attribution of movements may put in doubt the urgency for accurate absolute tracking in a range of full body applications in consumer VR. Notably, we argue that the main priority of motion sensors are neither precision in terms of absolute position nor preserving movement dynamics, instead it is to accurately track the proximity between limbs/end effectors, thus preserving visuo-tactile stimuli on situations where self-contact is present.

6.3 Limitations and Future Work

The manipulation of viewpoint, posture and movement to guide successful interaction and its relation with the sense of embodiment is complex, and we only scratch the surface of the range of possibilities. This project will resume, taking special attention to movement distortion, but we also make considerations on the subject of viewpoint.

To begin, the inter-subject variance of the just noticeable difference is considerable. In a real life application, the ideal would be to personalize this setting. We plan to investigate this issue in two fronts: (1) with further analysis of the collected motion capture data, using machine learning algorithms and defining movement features, such as target over/under shooting; (2) with an experiment to identify whether electroencephalography (EEG) signal correlates with the self-attribution of distorted movements (brief description available in A).

Furthermore, the distortion model presented in Chapter 5 manipulates the amplitude of the movement required to complete a reaching. While we have shown that this affects the difficulty of the task, a second elementary way to manipulate difficulty is by altering the size of the target, making its (motor) interaction space bigger or smaller than its visual size. We have a complementary experiment on this subject, which is briefly presented in Appendix B.

Regarding the point of view, our results diverge from other comparative experiments. It remains unclear if this difference relates to higher order processes that provide a judgement of ownership, influenced by agency and the engagement in an involving task [Synofzik et al., 2008a], or if it is a product of the sensorimotor contingencies. A compelling new measurement that could clarify whether action in 3PP can augment ownership and self-location are those assessing the peripersonal space. Notably, [Noel et al., 2015] has shown that the classical 3PP full body illusion [Lenggenhager et al., 2007] results in a spatial drift of the peripersonal space with relation to the subjects body. More specifically, the boundaries of the peripersonal space are projected forward, toward the seen body. This protocol could help to disentangle the contribution of sensorimotor and task involvement to the sense of embodiment.

Moreover, one of the main practical limitations of 3PP is the virtual body occlusion. Body occlusion caused discomfort to some subjects when they had to walk in the virtual environment, as well as weaker performance when the subject had to interact with objects located in front of the virtual body. One way we have identifyied to address these limitations is by making the body semi-transparent. Indeed, [Martini et al., 2015] present an experiment showing that one may develop the sense of body ownership of a semi-transparent virtual body seen in 1PP. In future work we would like to evaluate whether [Martini et al., 2015] findings extend to 3PP. In particular, we believe that the alternating perspective method that we propose could be enhanced by making the virtual body semi-transparent, but only when it is seen from a 3PP.

Nevertheless, we propose the use of redirected movement to manipulate the interaction of the virtual body with the virtual environment, which represents a broad range of possible movements and interaction strategies. However, we have only explored a small set of interactions in our experiments, i.e. goal directed reaching movements. We believe that the scenarios and situations where this approach is desirable should be better defined through experimentation.

Additionally, we have only manipulated movement based on a displacement of an end point. To contemplate a more complete context of interaction with the virtual environment, more dimensions of the movement and possible distortion should be considered. For instance, how do users perceive a movement that deviates from an obstacle in between start and end position of a movement? In this case the distortion would be orthogonal to the intended direction of movement, but would not interfere with the final position (except by inducing subjects to perform unnecessary corrective movements)

We expect to expand the subject of movement distortion in two main directions. By extending the complexity of distortion to account for a trajectory (not only a target position), and to better understand the relation of JND with the sense of ownership and agency of a full body. The former is especially relevant for virtual body interactions with the environment, as the trajectory could better account for environment constraints or obstacles. The latter involves transferring our current results on distortion to whole body interaction experiments.

Finally, we explored relatively new aspects of embodied interaction in virtual reality. Altogether we verified that visuo-motor contingencies of the whole body and meaningful interaction with the virtual environment have a significant influence on aspects of the sense of embodiment, and that subjects are receptive to movement distortions that facilitate the completion of a task. Despite only covering a limited range of sensorimotor manipulations and tasks, our findings add useful new guidelines to the VR interaction toolbox.

A Appendix

A.1 Neural signatures of self-movement and movements distortion in embodied VR

We seek to identify if movements above and below the JND threshold elicit known error associated waveform components of EEG ERPs. In particular, we look for traces of event related negativity (ERN), N400 (negative 400ms) and error positivity components (Pe).

- ERN is a short negative component starting $\approx 150ms$ after a visual stimuli. It is stronger at the fronto-central regions of the scalp, and is associated to a mismatch between expected and presented stimulus [Falkenstein et al., 2000].
- Pe is a long positive component starting $\approx 200ms$ after the visual stimuli. It is stronger at centro-parietal regions and is associated to commited errors [Falkenstein et al., 2000].
- N400 is a weaker negative component starting $\approx 400ms$ after the visual stimuli. It is stronger at the fronto-central regions and is associated to semantic incongruence [Kutas and Federmeier, 2011] and the observation of erroneous actions [Amoruso et al., 2013].

In the context of embodied VR, we highlight two experiments that explore these signals:

In the first, [Pavone et al., 2016] has performed an experiment in which subjects observe a virtual body performing a reaching (pre-recorded) movement toward a mug. The image is set so that the arm seems to stretch out of the subject shoulders. They demonstrate the three components could be observed when the virtual hand misses the mug. Notably, amplitude of ERN was associated with the sense of ownership of the virtual body.

In the second, [Padrao et al., 2016] performs an experiment in which a virtual hand could move to the wrong direction (relative to the instructed direction) either because of a system intervention or a mistake made by the subject. They report a stronger ERN specifically when the subject was the one to cause the incongruent event, and a stronger N400 when the system manipulated the movement.

We use the same setup presented in Chapter 5. The subject performs a reaching movements towards the right, and a small subset of distortions are used, comprising correct movement, a distortion below JND, and a distortion above JND. We hypothesize that an ERN may be found below the JND threshold, as the subject might self-attribute errors and required adjustments at this range, and that an N400 may be found above the JND threshold, as the source of the error might be attributed to an external source. However, we highlight the fact that the error we present is relatively more subtle than those of related work, and it may be too subtle to actually produce these components at a recognizable amplitude. Finally, we also experiment with machine learning algorithms to to use the ERP waveform as predictor to the subject answer.

B Appendix

B.1 Manipulating movement precision

In this experiment we manipulate the difficulty to accurately complete a tapping task. The tapping task could be made easier or harder by changing the mapping from the real to the virtual hand, i.e. the virtual hand position could diverge from the real hand position. More specifically, the interaction (motor) area of the target could be made bigger or smaller than its visual size. Thus, we warp the space around the target, being capable of fitting a bigger or smaller physical area than the virtual (visual) feedback suggests, effectively facilitating or hindering the completion of the task.

The subject had to perform a multi-directional pointing task, as described by the annex B of ISO 9241-411 [ISO, 2012]. This task consists of multiple pointing movements – 11 in this case – toward targets equally spaced over the borders of a circle. Subsequent targets are defined as to maximize the distance between them.

After each round of movements we make two questions:

- Did the virtual hand moved like you?
- Did you missed any target?

The first question is meant to assess the JND of the proposed manipulation (more detains on JND in Section 5.4). The second question is meant to disentangle the subjective bias of being unsuccessful in a given task from being capable of detecting a distortion. An overview of a pointing trial is shown in Figure B.1.

B.1.1 Distortion Model

The tapping task could be made easier or harder by manipulating the position mapping of the real to the virtual hand ($\mathbf{p}_{real}$ and $\mathbf{p}_{virtual}$ respectively) i.e. the virtual and real hands

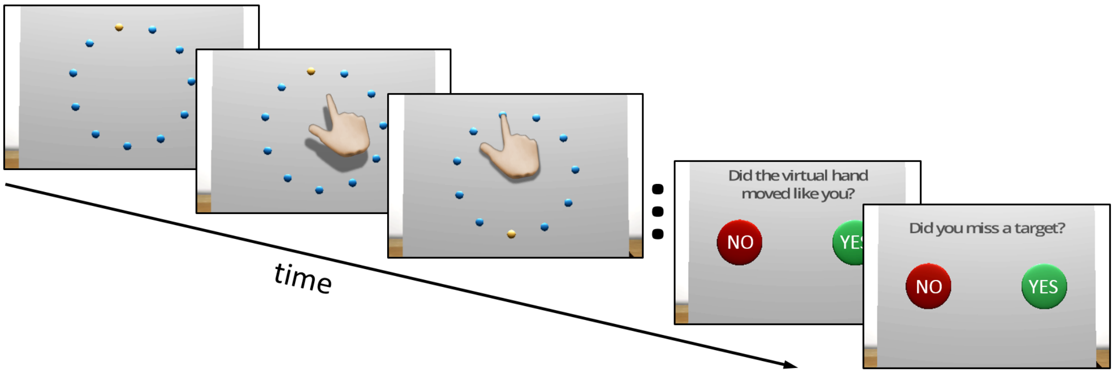


Figure B.1 – Overview of the pointing experiment trial. The subject had to tap 11 targets in a multi-direction pointing task. The current target is highlighted in orange. The movement could be distorted in the region surrounding the each target. Once the subjects tap the last target, they are asked whether the seen movement corresponds to the performed, and whether they have missed a target.

positions could diverge.

A distance range (d_{range}) is used so that when the target to hand distance (d_{real}) is bigger than d_{range} no remapping occurs. The d_{range} is also used to normalize the values to the range $[0, 1]$, and then scale this normalized remapping back into world units.

Our dynamic remapping uses properties of exponentiation of values between 0 and 1 so that when the exponent is bigger than 1 the interaction width of the target becomes bigger, and when it is less than 1 the interaction width of the target is reduced. Thus facilitating and hindering the completion of the goal directed task (Figure B.2).

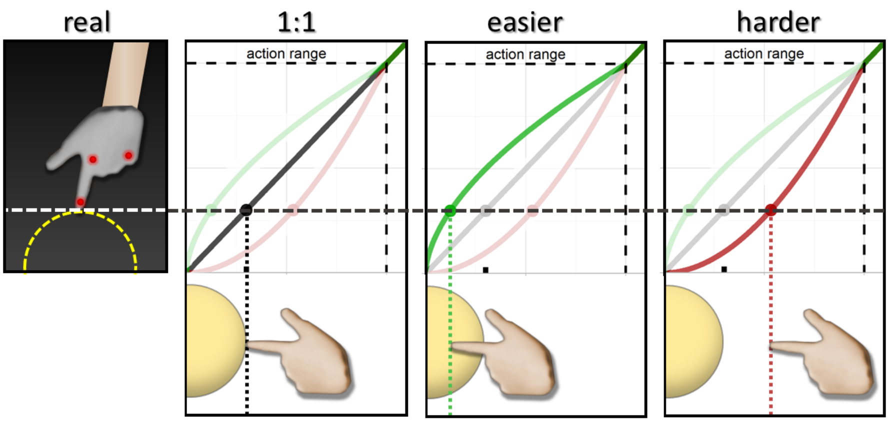


Figure B.2 – Overview of the pointing distortion function. The real movement (left) is mapped into a virtual movement according to the settings of the function (right). The lower left corner of the graphics represents the center of the target.

The distortion method is presented below:

$$d_{virtual} = \begin{cases} d_{range} \times (\frac{d_{real}}{d_{range}})^a, & \text{if } d_{real} < d_{range} \\ d_{real}, & \text{otherwise} \end{cases} \tag{B.1}$$

Where $d_{real} = ||v_{dir}||$ and $v_{dir} = p_{real} - p_{tgt}$. The exponent a is defined by

$$a = \log_{\frac{tgt_{interaction}}{d_{range}}}(\frac{tgt}{d_{range}})$$

Where tgt represents the radius of the target, and $tgt_{interaction}$ represents a *desired* radius of the interaction target. Therefore, when $tgt_{interaction} < tgt$ the task becomes harder, and when $tgt_{interaction} > tgt$ the task becomes easier. Mind that our distortion model assumes that tgt and $tgt_{interaction}$ are smaller than d_{range}.

The current position of the virtual hand ($p_{virtual}$) is then set using:

$$\mathbf{p}_{virtual} = \mathbf{p}_{tgt} - \hat{\mathbf{v}_{dir}} \times d_{virtual}$$

where $\mathbf{v}_{dir} = \mathbf{p}_{real} - \mathbf{p}_{tgt}$

B.1.2 Materials and Methods

Equipment

The participant wore an Oculus Development Kit 2 HMD to visualize the virtual scene. A PhaseSpace ImpulseX2 system with 18 cameras was used to capture participant and virtual object movements. A total of 14 LED markers were used, 4 attached to the HMD, 4 attached to the hand, 3 attached to the table, and 4 attached to the tapping surface. The glove had a marker over the index fingertip, and two over the back of the hand. A rigid and flat stick was positioned between the top of the subject's index finger and the glove in order to prevent the finger from flexing. The table and the tapping surface were also tracked. Figure B.3 presents an overview of the setup.

The markers on the glove were pre-calibrated. To compensate for small changes in length of the index finger, we calibrate the tapping surface. The calibration consisted of pointing at three predefined corners of the tapping surface. The plane defined by these corners was then used to translate and rotate the tapping surface into a compatible position.

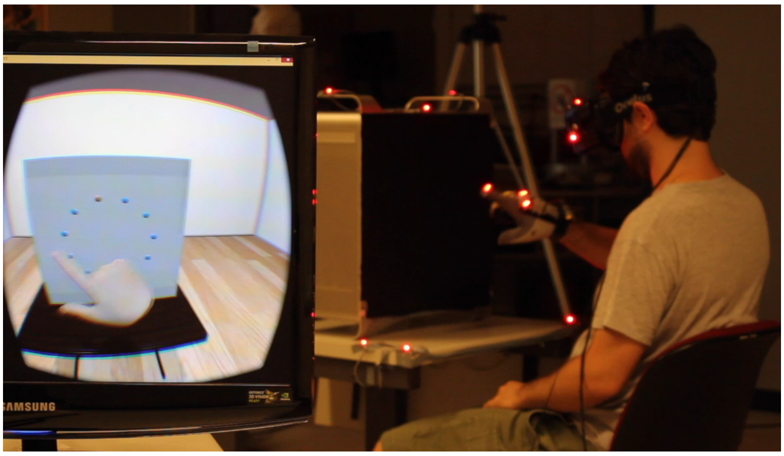


Figure B.3 – Overview of the pointing experiment setup. The subject sit in a chair and wore an HMD and a motion tracking glove.

Experiment design

We manipulated two variables, the visual index of difficulty of the task (*visualID*), and the difference in the index of difficulty caused by the distortion (*diffID*). Details on the index of difficulty and its relation with distance of movement and size of target in a reaching task were presented in Chapter 5. The *visualID* could be set to 4 or 5, we did so while keeping the distance constant ($27cm$) and solving for the required width of the target using the equation $tgt = \frac{D}{2^{visualID}-1}$ where $D = distance$. The *diffID* could be set to -2.5, -2, -1.5, -1, -0.5, 0 (no distortion), 0.5, 1, 1.5, 2 or 2.5. A positive *diffID* means that the motor size of the target became smaller than the visual feedback suggests (i.e. harder), while a negative *diffID* had the opposite effect (i.e. easier). Similarly to tgt, the $tgt_{interaction}$ is computed with the equation $tgt_{interaction} = \frac{D}{2^{visualID+diffID}-1}$.

We design this experiment to analyze two aspects of redirected interaction:

First, we want to assess the **just noticeable difference (JND)** for this distortion model and task, i.e. **the thresholds after which the distortion becomes likely to be perceived**.

Second, we want to replicate the bias presented in Chapter 5, i.e. confirm that **subjects are biased to self-attribute movement distortions that make the task easier**. We can confirm that if we obtain a negative **point of subjective equality (PSE)** in the *diffID* and if subjects self-attribute movements when they report no missed target ("no" to second question) more often than when they report to have missed a target ("yes" to second question).

B.1.3 Results and Conclusions

We received 15 subjects, two of whom have been excluded due to very poor distortion recognition performance.

The JNDs and the PSE were obtained by fitting a normal distribution to the data of each subject. The PSE is equivalent to the mean of the distribution, while the JNDs were computed as the PSE$\pm 1.178 \times SD$.

We found the following JNDs thresholds (in $diffID$ units) $1.20 \pm .56$ and $-1.73 \pm .74$ for the hindering and facilitating distortions respectively. Thus, distortions within this range of $diffID$ values were more likely to be self-attributed than not (Figure B.4).

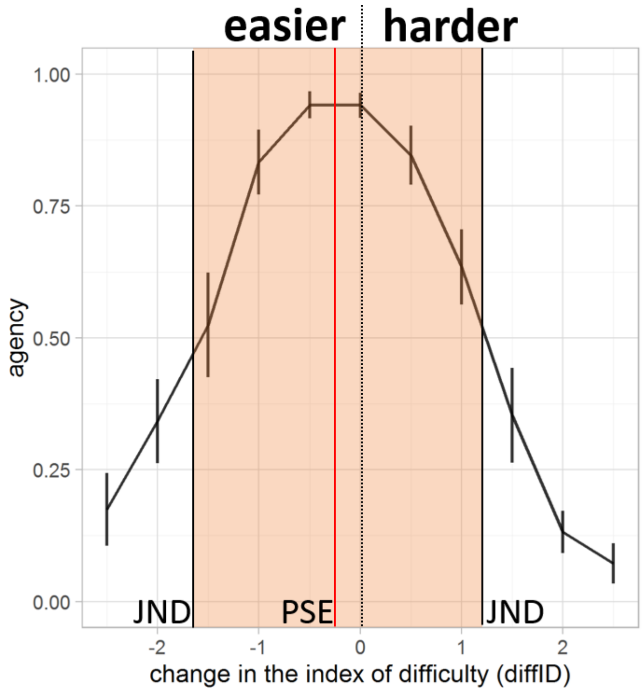


Figure B.4 – Pointing experiment JND and PSE results. The graphic presents the mean self-attribution for the different levels of *diffID*. The vertical lines represent the mean JNDs and PSE found by fitting a normal distribution to the data of each subject. Subjects tended to self-attribute distorted movements more often when the distortion made the task slightly easier than the unmodified movement. Error bars represent the standard error of the mean.

We obtained a meanpmSD of $-.27 \pm .39$, a t-test shows that the PSE was significantly smaller than zero ($t_{12} = 2.44$ $p < .032$, Figure B.4). Moreover, subject were less likely to self-attribute a movement when they were aware that a at least one target in the trial had been missed ($t_{12} = 8.36$ $p < .001$, Figure B.5). These results support the notion that subjects are biased toward self-attributing movements that make the task easier to complete, as suggested in Chapter 5.

In this experiment we expand the characteristics of the movement distortion model by manipulating the size of the target, instead of the distance to reach it (as in Chapter 5). We show that the bias toward self-attribution of movements that make the task easier is also present for this family of manipulation, therefore making the findings of Chapter 5 more consistent.

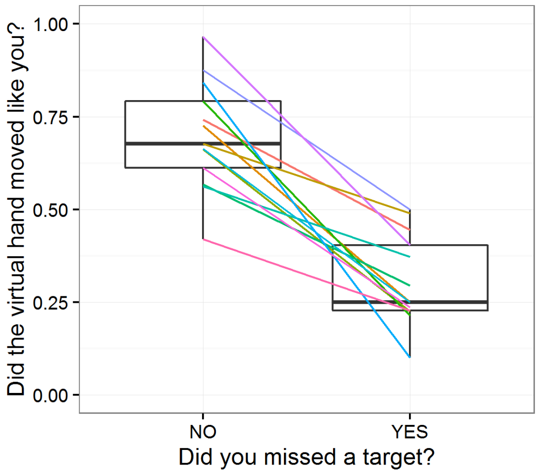


Figure B.5 – Pointing experiment relation of self-attribution and perception of errors. Subjects were more likely to self-attribute a movements when they were not aware of pointing errors.

C Appendix

Additional materials for the study: **Effect of perspective and visuo-motor synchrony to the sense of embodiment** presented in Section 4.2

1. Pre-experiment characterization questionnaire.
2. Embodiment questionnaire applied after each experimental condition.

Edit this form

CHARACTERIZATION

All the gathered data will be treated anonymously.

* Required

Identifier (filled by the experimenter) *

Height *

Weight *

Age *

Gender *

- Male
- Female

How often do you participate on experiments using Virtual Reality equipments? *

e.g. experiments in other labs of the university.

- Never participated of an experiment
- A few times
- Every month
- Every week
- Every day

How often do you use head mounted displays? *

- Never used
- A few times
- Every month
- Every week
- Every day

How often do you play video games? *

- Never played
- A few times
- Every month
- Every week
- Every day

How often do use the Microsoft Kinect, Nintendo Wii or Playstation move? *

- Never used
- A few times
- Every month
- Every week
- Every day

Hand of preference *

usually, the hand you write with

- Left hand
- Right hand

Area(s) of expertise/study/work/interest *

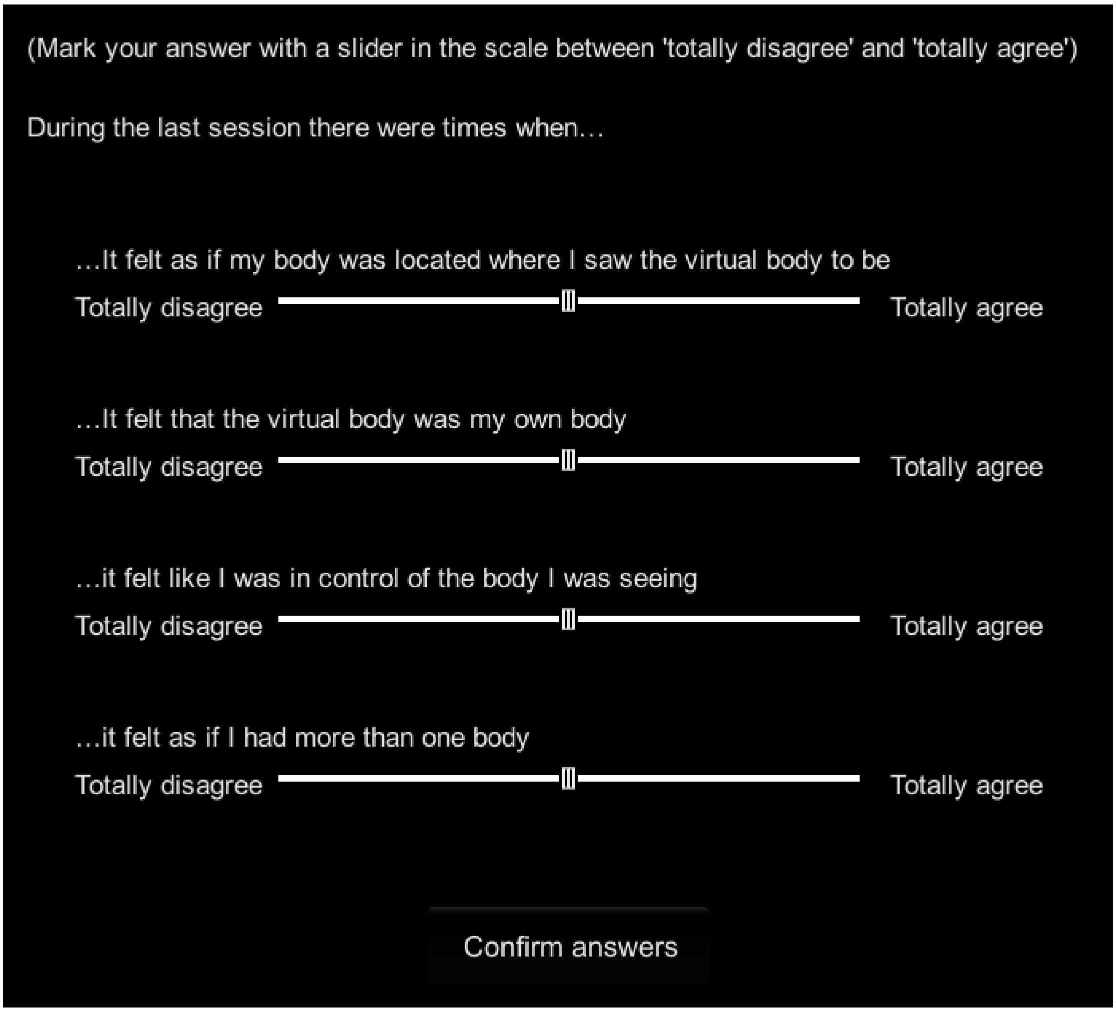
e.g. computer science, math, sales, mechanical engineering etc.

Are you a student? *

- Yes, bachelor student
- Yes, master student
- Yes, PhD student
- No, I'm not a student

Submit

Never submit passwords through Google Forms.


(Mark your answer with a slider in the scale between 'totally disagree' and 'totally agree')

During the last session there were times when…

…It felt as if my body was located where I saw the virtual body to be

Totally disagree — Totally agree

…It felt that the virtual body was my own body

Totally disagree — Totally agree

…it felt like I was in control of the body I was seeing

Totally disagree — Totally agree

…it felt as if I had more than one body

Totally disagree — Totally agree

Confirm answers

D Appendix

Additional materials for the study: **Effect of perspective and multi-modal congruence to the sense of embodiment** presented in Section 4.3

Pre-experiment characterization questionnaire were the same presented in Appendix C

1. Embodiment questionnaire applied after each experimental condition.
2. Post-experiment perspective questionnaire.

Edit this form

Questionnaire

* Required

Subject ID *
(filled by the experimenter)

Condition *
(filled by the experimenter)

- First (1) Person Perspective
- Third (3) Person Perspective
- First/Third (1/3) Person Perspective

Read carefully

During the last session ...

... it felt as if my body was located where I saw the virtual body to be *

-3 strongly DISAGREE	-2	-1	0	1	2	3 strongly AGREE

... I felt as if I was looking to my own body *

-3 strongly DISAGREE	-2	-1	0	1	2	3 strongly AGREE

... It seemed as if I were sensing the movement of my body in the location where the virtual body moved *

-3 strongly DISAGREE	-2	-1	0	1	2	3 strongly AGREE

... it felt as if I could get hurt if the virtual body was to fall in the pit *

-3 strongly DISAGREE	-2	-1	0	1	2	3 strongly AGREE

○	○	○	○	○	○	○

...It felt that the virtual body was my own body *

-3 strongly DISAGREE	-2	-1	0	1	2	3 strongly AGREE
○	○	○	○	○	○	○

... whenever I moved my body I expected the virtual body to move in the same way *

-3 strongly DISAGREE	-2	-1	0	1	2	3 strongly AGREE
○	○	○	○	○	○	○

...it felt as if my real body was turning virtual *

-3 strongly DISAGREE	-2	-1	0	1	2	3 strongly AGREE
○	○	○	○	○	○	○

...it felt as if I had more than one body *

-3 strongly DISAGREE	-2	-1	0	1	2	3 strongly AGREE
○	○	○	○	○	○	○

...it felt like I was in control of the body I was seeing *

-3 strongly DISGREE	-2	-1	0	1	2	3 strongly AGREE
○	○	○	○	○	○	○

... I felt as if the pit posed a threat to myself *

-3 strongly DISAGREE	-2	-1	0	1	2	3 strongly AGREE
○	○	○	○	○	○	○

Submit

Never submit passwords through Google Forms.

Edit this form

POST EXPERIMENT

* Required

Subject ID *
(filled by the experimenter)

Which point of view do you prefer to use to walk forward? *

- THIRD person perspective
- FIRST person perspective

Which point of view makes you feel safer when the floor falls? *

- THIRD person perspective
- FIRST person perspective

Which point of view do you prefer to use when the floor falls? *

- THIRD person perspective
- FIRST person perspective

Which point of view do you prefer to use to reach the targets? *

- FIRST person perspective
- THIRD person perspective

Which condition do you PREFER to perform the reaching task? *

- THIRD person perspective alone
- Being able to switch between FIRST and THIRD person perspective
- FIRST person perspective alone

Which condition do you think to be MORE EFFICIENT to perform the reaching task? *

- FIRST person perspective alone
- Being able to switch between FIRST and THIRD person perspective
- THIRD person perspective alone

Which floor did you consider when performing the Mental Ball Drop? *

- The floor of the virtual environment
- The floor of the lab

Submit

Never submit passwords through Google Forms.

E Appendix

Additional materials for the study: **Embodied interaction and non-planar projections in immersive virtual reality** presented in Section 3.1

Pre-experiment characterization questionnaire were the same presented in Appendix C

1. Embodiment questionnaire applied after each experimental condition.
2. Simulation Sickness questionnaire (provided by the Cyberpsychology Lab of the University of Quebec).

Embodiment Questionnaire

1. **Subject ID**

....

Read carefully

(Mark your answer with a cross in the scale between “strongly disagree” and “strongly agree”)

During the last session there were times when...

2. **...it felt like the body I was seeing was out of my control** *

Une seule réponse possible.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

3. **...it felt as if I had more than one body** *

Une seule réponse possible.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

4. **...it felt like I could move the body I was seeing as I wanted** *

Une seule réponse possible.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

5. **...it felt as if my real body was turning virtual** *

Une seule réponse possible.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

6. **...it felt like I was in control of the body I was seeing** *

Une seule réponse possible.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

7. **... I felt as if I were wearing the clothes I was seeing, rather than my actual clothes** *

Une seule réponse possible.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

8. **... it felt as if I could really touch the shelves I was seeing** *

Une seule réponse possible.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

9. **...it felt that the body I was seeing was not my body** *

Une seule réponse possible.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

10. **...It felt that the body I was seeing was my own body** *

Une seule réponse possible.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

No______________ Date____________________

SIMULATOR SICKNESS QUESTIONNAIRE

Kennedy, Lane, Berbaum, & Lilienthal (1993)***

Instructions : Circle how much each symptom below is affecting you right now.

1. General discomfort	None	Slight	Moderate	Severe
2. Fatigue	None	Slight	Moderate	Severe
3. Headache	None	Slight	Moderate	Severe
4. Eye strain	None	Slight	Moderate	Severe
5. Difficulty focusing	None	Slight	Moderate	Severe
6. Salivation increasing	None	Slight	Moderate	Severe
7. Sweating	None	Slight	Moderate	Severe
8. Nausea	None	Slight	Moderate	Severe
9. Difficulty concentrating	None	Slight	Moderate	Severe
10. « Fullness of the Head »	None	Slight	Moderate	Severe
11. Blurred vision	None	Slight	Moderate	Severe
12. Dizziness with eyes open	None	Slight	Moderate	Severe
13. Dizziness with eyes closed	None	Slight	Moderate	Severe
14. *Vertigo	None	Slight	Moderate	Severe
15. **Stomach awareness	None	Slight	Moderate	Severe
16. Burping	None	Slight	Moderate	Severe

* Vertigo is experienced as loss of orientation with respect to vertical upright.

** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

Last version : March 2013

***Original version : Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3*(3), 203-220.

Bibliography

[Amoruso et al., 2013] Amoruso, L., Gelormini, C., Aboitiz, F. A., González, A., Manes, F., Cardona, J., Ibanez, A., et al. (2013). N400 erps for actions: building meaning in context. *Frontiers in human neuroscience*, 7:57.

[Ardouin et al., 2013] Ardouin, J., Lécuyer, A., Marchal, M., and Marchand, E. (2013). Navigating in virtual environments with 360° omnidirectional rendering. In *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, pages 95–98. IEEE.

[Ardouin et al., 2012] Ardouin, J., Lécuyer, A., Marchal, M., Riant, C., and Marchand, E. (2012). Flyviz: a novel display device to provide humans with 360 vision by coupling catadioptric camera with hmd. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*, pages 41–44. ACM.

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

[Aspell et al., 2009] Aspell, J. E., Lenggenhager, B., and Blanke, O. (2009). Keeping in touch with one's self: multisensory mechanisms of self-consciousness. *PloS one*, 4(8):e6488.

[Bailenson et al., 2008] Bailenson, J., Patel, K., Nielsen, A., Bajscy, R., Jung, S.-H., and Kurillo, G. (2008). The effect of interactivity on learning physical actions in virtual reality. *Media Psychology*, 11(3):354–376.

[Ban et al., 2012a] Ban, Y., Kajinami, T., Narumi, T., Tanikawa, T., and Hirose, M. (2012a). Modifying an identified curved surface shape using pseudo-haptic effect. In *Haptics Symposium (HAPTICS), 2012 IEEE*, pages 211–216. IEEE.

[Ban et al., 2012b] Ban, Y., Narumi, T., Tanikawa, T., and Hirose, M. (2012b). Modifying an identified position of edged shapes using pseudo-haptic effects. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*, pages 93–96. ACM.

[Ban et al., 2014] Ban, Y., Narumi, T., Tanikawa, T., and Hirose, M. (2014). Displaying shapes with various types of surfaces using visuo-haptic interaction. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pages 191–196. ACM.

[Banakou and Slater, 2014] Banakou, D. and Slater, M. (2014). Body ownership causes illusory self-attribution of speaking and influences subsequent real speaking. *Proceedings of the National Academy of Sciences*, 111(49):17678–17683.

[Bernstein, 2013] Bernstein, D. (2013). *Essentials of psychology*. Cengage Learning.

[Biocca, 1997] Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual environments [1]. *Journal of Computer-Mediated Communication*, 3(2):0–0.

[Blakemore et al., 1999] Blakemore, S.-J., Frith, C. D., and Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of cognitive neuroscience*, 11(5):551–559.

[Blakemore et al., 2002] Blakemore, S.-J., Wolpert, D. M., and Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in cognitive sciences*, 6(6):237–242.

[Blanke, 2012] Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*, 13(8):556–571.

[Blanke and Metzinger, 2009] Blanke, O. and Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends in cognitive sciences*, 13(1):7–13.

[Blanke et al., 2015] Blanke, O., Slater, M., and Serino, A. (2015). Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron*, 88(1):145–166.

[Botvinick, 2004] Botvinick, M. (2004). NEUROSCIENCE: probing the neural basis of body ownership. *Science*, 305(5685):782–783.

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

[Boulic et al., 2010] Boulic, R., Maupu, D., Peinado, M., and Raunhardt, D. (2010). Spatial awareness in full-body immersive interactions: where do we stand? In *Motion in Games*, pages 59–69. Springer.

[Boulic et al., 2009] Boulic, R., Maupu, D., and Thalmann, D. (2009). On scaling strategies for the full-body postural control of virtual mannequins. *Interacting With Computers*, 21:11–25.

[Bourke, 2009] Bourke, P. (2009). idome: Immersive gaming with the unity game engine. In *Proceedings of the Computer Games & Allied Technology*, pages 265–272.

[Bowman et al., 1997] Bowman, D. A., Koller, D., and Hodges, L. F. (1997). Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *Virtual Reality Annual International Symposium, 1997., IEEE 1997*, pages 45–52. IEEE.

[Bowman and McMahan, 2007] Bowman, D. A. and McMahan, R. P. (2007). Virtual reality: how much immersion is enough? *Computer*, 40(7):36–43.

[Box and Cox, 1964] Box, G. E. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 211–252.

[Brondi et al., 2015] Brondi, R., Alem, L., Avveduto, G., Faita, C., Carrozzino, M., Tecchia, F., and Bergamasco, M. (2015). Evaluating the impact of highly immersive technologies and natural interaction on player engagement and flow experience in games. In *International Conference on Entertainment Computing*, pages 169–181. Springer.

[Burns and Brooks, 2006] Burns, E. and Brooks, F. P. (2006). Perceptual sensitivity to visual/kinesthetic discrepancy in hand speed, and why we might care. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '06, pages 3–8, New York, NY, USA. ACM.

[Burns et al., 2006] Burns, E., Razzaque, S., Panter, A. T., Whitton, M. C., McCallus, M. R., and Brooks Jr, F. P. (2006). The hand is more easily fooled than the eye: Users are more sensitive to visual interpenetration than to visual-proprioceptive discrepancy. *Presence: Teleoperators and Virtual Environments*, 15(1):1–15.

[Burns et al., 2007] Burns, E., Razzaque, S., and Whitton, M. C. (2007). Macbeth: Management of avatar conflict by employment of a technique hybrid. *International Journal of Virtual Reality*, 68(04):11–20.

[Caspar et al., 2015] Caspar, E. A., Cleeremans, A., and Haggard, P. (2015). The relationship between human agency and embodiment. *Consciousness and cognition*, 33:226–236.

[Cleveland and Devlin, 1988] Cleveland, W. S. and Devlin, S. J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403):596–610.

[Covaci et al., 2014] Covaci, A., Olivier, A.-H., and Multon, F. (2014). Third person view and guidance for more natural motor behaviour in immersive basketball playing. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, VRST '14, pages 55–64, New York, NY, USA. ACM.

[Csikszentmihalyi and Csikszentmihalyi, 1992] Csikszentmihalyi, M. and Csikszentmihalyi, I. S. (1992). *Optimal experience: Psychological studies of flow in consciousness*. Cambridge university press.

[David et al., 2008] David, N., Newen, A., and Vogeley, K. (2008). The "sense of agency" and its underlying cognitive and neural mechanisms. *Consciousness and cognition*, 17(2):523–534.

[De Vignemont, 2011] De Vignemont, F. (2011). Embodiment, ownership and disownership. *Consciousness and cognition*, 20(1):82–93.

[De Vignemont and Fourneret, 2004] De Vignemont, F. and Fourneret, P. (2004). The sense of agency: A philosophical and empirical review of the "who" system. *Consciousness and Cognition*, 13(1):1–19.

[Debarba et al., 2013] Debarba, H. G., Molla, E., Herbelin, B., and Boulic, R. (2013). 4pp: augmenting 1pp with 3pp. In *Joint Virtual Reality Conference of EGVE-EuroVR JVRC 2013*, pages 83–86. EuroVR.

[Debarba et al., 2015a] Debarba, H. G., Molla, E., Herbelin, B., and Boulic, R. (2015a). Characterizing embodied interaction in first and third person perspective viewpoints. In *3D User Interfaces (3DUI), 2015 IEEE Symposium on*, pages 67–72. IEEE.

[Debarba et al., 2015b] Debarba, H. G., Perrin, S., Herbelin, B., and Boulic, R. (2015b). Embodied interaction using non-planar projections in immersive virtual reality. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, pages 125–128. ACM.

[Desmurget et al., 1999] Desmurget, M., Epstein, C., Turner, R., Prablanc, C., Alexander, G., and Grafton, S. (1999). Role of the posterior parietal cortex in updating reaching movements to a visual target. *Nature neuroscience*, 2(6):563–567.

[Desmurget and Grafton, 2000] Desmurget, M. and Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in cognitive sciences*, 4(11):423–431.

[Falkenstein et al., 2000] Falkenstein, M., Hoormann, J., Christ, S., and Hohnsbein, J. (2000). Erp components on reaction errors and their functional significance: a tutorial. *Biological psychology*, 51(2):87–107.

[Franck et al., 2001] Franck, N., Farrer, C., Georgieff, N., Marie-Cardine, M., Daléry, J., d'Amato, T., and Jeannerod, M. (2001). Defective recognition of one's own actions in patients with schizophrenia. *American Journal of Psychiatry*, 158(3):454–459.

[Frith, 2012] Frith, C. (2012). Explaining delusions of control: The comparator model 20years on. *Consciousness and cognition*, 21(1):52–54.

[Frith et al., 2000] Frith, C. D., Wolpert, D. M., et al. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 355(1404):1771–1788.

[Gallagher, 2000] Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in cognitive sciences*, 4(1):14–21.

[Glaeser and Gröller, 1999] Glaeser, G. and Gröller, E. (1999). Fast generation of curved perspectives for ultra-wide-angle lenses in vr applications. *The Visual Computer*, 15(7-8):365–376.

[González-Franco et al., 2010] González-Franco, M., Pérez-Marcos, D., Spanlang, B., and Slater, M. (2010). The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *Virtual Reality Conference (VR), 2010 IEEE*, pages 111–114. IEEE.

[Haggard and Chambon, 2012] Haggard, P. and Chambon, V. (2012). Sense of agency. *Current Biology*, 22(10):R390–R392.

[Havranek et al., 2012] Havranek, M., Langer, N., Cheetham, M., and Jäncke, L. (2012). Perspective and agency during video gaming influences spatial presence experience and brain activation patterns. *Behavioral and Brain Functions*, 8(1):34.

[Heeter, 1992] Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators & Virtual Environments*, 1(2):262–271.

[Held and Durlach, 1992] Held, R. M. and Durlach, N. I. (1992). Telepresence. *Presence: Teleoperators & Virtual Environments*, 1(1):109–112.

[Hirose et al., 2006] Hirose, K., Ogawa, T., Kiyokawa, K., and Takemura, H. (2006). Interactive reconfiguration techniques of reference frame hierarchy in the multi-viewport interface. In *3D User Interfaces, 2006. 3DUI 2006. IEEE Symposium on*, pages 73–80. IEEE.

[ISO, 2012] ISO (2012). Ergonomics of human-system interaction – part 411: Evaluation methods for the design of physical input devices. ISO 9241-411:2012, International Organization for Standardization, Geneva, Switzerland.

[Jauregui et al., 2014] Jauregui, D. A. G., Argelaguet, F., Olivier, A.-H., Marchal, M., Multon, F., and Lecuyer, A. (2014). Toward" pseudo-haptic avatars": Modifying the visual animation of self-avatar can simulate the perception of weight lifting. *IEEE transactions on visualization and computer graphics*, 20(4):654–661.

[Jeannerod, 2003] Jeannerod, M. (2003). The mechanism of self-recognition in humans. *Behavioural brain research*, 142(1):1–15.

[Jeannerod, 2004] Jeannerod, M. (2004). Visual and action cues contribute to the self–other distinction. *Nature neuroscience*, 7(5):422–423.

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

[Kalckert and Ehrsson, 2012b] Kalckert, A. and Ehrsson, H. H. (2012b). Moving a rubber hand that feels like your own: A dissociation of ownership and agency. *Frontiers in Human Neuroscience*, 6.

[Kannape and Blanke, 2012] Kannape, O. and Blanke, O. (2012). Agency, gait and self-consciousness. *International Journal of Psychophysiology*, 83(2):191–199.

[Kannape and Blanke, 2013] Kannape, O. A. and Blanke, O. (2013). Self in motion: sensorimotor and cognitive mechanisms in gait agency. *Journal of neurophysiology*, 110(8):1837–1847.

[Kawato, 1999] Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current opinion in neurobiology*, 9(6):718–727.

[Kennedy and Koop, 1994] Kennedy, M. and Koop, S. (1994). Understanding map projections. *GIS by ESRI, Environmental System Research Institute.*

[Kennedy et al., 1993] Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220.

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

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

[Kiyokawa and Takemura, 2005] Kiyokawa, K. and Takemura, H. (2005). A tunnel window and its variations: Seamless teleportation techniques in a virtual environment. In *HCI International*.

[Kohli, 2010] Kohli, L. (2010). Redirected touching: Warping space to remap passive haptics. In *3D User Interfaces (3DUI), 2010 IEEE Symposium on*, pages 129–130. IEEE.

[Kohli et al., 2012] Kohli, L., Whitton, M. C., and Brooks, F. P. (2012). Redirected touching: The effect of warping space on task performance. In *3D User Interfaces (3DUI), 2012 IEEE Symposium on*, pages 105–112. IEEE.

[Kohli et al., 2013] Kohli, L., Whitton, M. C., and Brooks, F. P. (2013). Redirected touching: Training and adaptation in warped virtual spaces. In *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, pages 79–86. IEEE.

[Kokkinara et al., 2016] Kokkinara, E., Kilteni, K., Blom, K. J., and Slater, M. (2016). First person perspective of seated participants over a walking virtual body leads to illusory agency over the walking. *Scientific Reports*, 6(28879).

[Kokkinara and Slater, 2014] Kokkinara, E. and Slater, M. (2014). Measuring the effects through time of the influence of visuomotor and visuotactile synchronous stimulation on a virtual body ownership illusion. *Perception*, 43(1):43–58.

[Kokkinara et al., 2015] Kokkinara, E., Slater, M., and López-Moliner, J. (2015). The effects of visuomotor calibration to the perceived space and body, through embodiment in immersive virtual reality. *ACM Trans. Appl. Percept.*, 13(1):3:1–3:22.

[Kutas and Federmeier, 2011] Kutas, M. and Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the n400 component of the event related brain potential (erp). *Annual review of psychology*, 62:621.

[LaViola Jr, 2000] LaViola Jr, J. J. (2000). A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin*, 32(1):47–56.

[Lécuyer et al., 2008] Lécuyer, A., Burkhardt, J.-M., and Tan, C.-H. (2008). A study of the modification of the speed and size of the cursor for simulating pseudo-haptic bumps and holes. *ACM Trans. Appl. Percept.*, 5(3):14:1–14:21.

[Lecuyer et al., 2000] Lecuyer, A., Coquillart, S., Kheddar, A., Richard, P., and Coiffet, P. (2000). Pseudo-haptic feedback: can isometric input devices simulate force feedback? In *Virtual Reality, 2000. Proceedings. IEEE*, pages 83–90. IEEE.

[Legrand, 2006] Legrand, D. (2006). The bodily self: The sensori-motor roots of pre-reflective self-consciousness. *Phenomenology and the Cognitive Sciences*, 5(1):89–118.

[Lenggenhager et al., 2009] Lenggenhager, B., Mouthon, M., and Blanke, O. (2009). Spatial aspects of bodily self-consciousness. *Consciousness and cognition*, 18(1):110–117.

[Lenggenhager et al., 2006] Lenggenhager, B., Smith, S. T., and Blanke, O. (2006). Functional and neural mechanisms of embodiment: importance of the vestibular system and the temporal parietal junction. *Reviews in the Neurosciences*, 17(6):643–657.

[Lenggenhager et al., 2007] Lenggenhager, B., Tadi, T., Metzinger, T., and Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. *Science*, 317(5841):1096–1099.

[Lombard and Ditton, 1997] Lombard, M. and Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2):0–0.

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

[Lopez et al., 2014] Lopez, T., Bouville, R., Loup-Escande, E., Nouviale, F., Gouranton, V., and Arnaldi, B. (2014). Exchange of avatars: Toward a better perception and understanding. *IEEE transactions on visualization and computer graphics*, 20(4):644–653.

[Martini et al., 2015] Martini, M., Kilteni, K., Maselli, A., and Sanchez-Vives, M. V. (2015). The body fades away: investigating the effects of transparency of an embodied virtual body on pain threshold and body ownership. *Scientific reports*, 5.

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

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

[Maupu et al., 2009] Maupu, D., Boulic, R., and Thalmann, D. (2009). Characterizing full-body reach duration across task and viewpoint modalities. *Journal of Virtual Reality and Broadcasting*, 5(15).

[Meehan et al., 2002] Meehan, M., Insko, B., Whitton, M., and Brooks Jr, F. P. (2002). Physiological measures of presence in stressful virtual environments. *ACM Transactions on Graphics (TOG)*, 21(3):645–652.

[Meese, 1995] Meese, T. (1995). Using the standard staircase to measure the point of subjective equality: A guide based on computer simulations. *Perception & Psychophysics*, 57(3):267–281.

[Molla and Boulic, 2013] Molla, E. and Boulic, R. (2013). Singularity free parametrization of human limbs. In *Proceedings of the Motion on Games*, pages 165–174. ACM.

[Mulloni et al., 2012] Mulloni, A., Seichter, H., Dünser, A., Baudisch, P., and Schmalstieg, D. (2012). 360 panoramic overviews for location-based services. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2565–2568. ACM.

[Nielsen, 1963] Nielsen, T. I. (1963). Volition: A new experimental approach. *Scandinavian journal of psychology*, 4(1):225–230.

[Noel et al., 2015] Noel, J.-P., Pfeiffer, C., Blanke, O., and Serino, A. (2015). Peripersonal space as the space of the bodily self. *Cognition*, 144:49–57.

[Normand et al., 2011] Normand, J.-M., Giannopoulos, E., Spanlang, B., and Slater, M. (2011). Multisensory stimulation can induce an illusion of larger belly size in immersive virtual reality. *PLoS ONE*, 6(1).

[Orlosky et al., 2014] Orlosky, J., Wu, Q., Kiyokawa, K., Takemura, H., and Nitschke, C. (2014). Fisheye vision: Peripheral spatial compression for improved field of view in head mounted displays. In *Proceedings of the 2Nd ACM Symposium on Spatial User Interaction*, SUI '14, pages 54–61, New York, NY, USA. ACM, ACM.

[Padrao et al., 2016] Padrao, G., Gonzalez-Franco, M., Sanchez-Vives, M. V., Slater, M., and Rodriguez-Fornells, A. (2016). Violating body movement semantics: Neural signatures of self-generated and external-generated errors. *NeuroImage*, 124:147–156.

[Pausch et al., 1993] Pausch, R., Shackelford, M. A., and Proffitt, D. (1993). A user study comparing head-mounted and stationary displays. In *Virtual Reality, 1993. Proceedings., IEEE 1993 Symposium on Research Frontiers in*, pages 41–45. IEEE.

[Pavone et al., 2016] Pavone, E. F., Tieri, G., Rizza, G., Tidoni, E., Grisoni, L., and Aglioti, S. M. (2016). Embodying others in immersive virtual reality: electro-cortical signatures of monitoring the errors in the actions of an avatar seen from a first-person perspective. *The Journal of Neuroscience*, 36(2):268–279.

[Petkova et al., 2011] Petkova, V., Khoshnevis, M., and Ehrsson, H. (2011). The perspective matters! multisensory integration in ego-centric reference frames determines full-body ownership. *Frontiers in psychology*, 2.

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

[Pomés and Slater, 2013] Pomés, A. and Slater, M. (2013). Drift and ownership toward a distant virtual body. *Frontiers in human neuroscience*, 7.

[Ramachandran et al., 1991] Ramachandran, V. S., Gregory, R. L., et al. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature*, 350(6320):699–702.

[Ramachandran and Rogers-Ramachandran, 1996] Ramachandran, V. S. and Rogers-Ramachandran, D. (1996). Synaesthesia in phantom limbs induced with mirrors. *Proceedings of the Royal Society of London B: Biological Sciences*, 263(1369):377–386.

[Razzaque et al., 2001] Razzaque, S., Kohn, Z., and Whitton, M. C. (2001). Redirected walking. In *Proceedings of EUROGRAPHICS*, volume 9, pages 105–106. Citeseer.

[Renard et al., 2010] Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., and Lécuyer, A. (2010). Openvibe: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence: teleoperators and virtual environments*, 19(1):35–53.

[Rohrer et al., 2002] Rohrer, B., Fasoli, S., Krebs, H. I., Hughes, R., Volpe, B., Frontera, W. R., Stein, J., and Hogan, N. (2002). Movement smoothness changes during stroke recovery. *The Journal of Neuroscience*, 22(18):8297–8304.

[Salamin et al., 2010] Salamin, P., Tadi, T., Blanke, O., Vexo, F., and Thalmann, D. (2010). Quantifying effects of exposure to the third and First-Person perspectives in Virtual-Reality-Based training. *Learning Technologies, IEEE Transactions on*, 3(3):272–276.

[Salamin et al., 2008] Salamin, P., Thalmann, D., and Vexo, F. (2008). Improved third-person perspective: a solution reducing occlusion of the 3pp? In *Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, VRCAI '08, pages 30:1–30:6, New York, NY, USA. ACM.

[Sanchez-Vives and Slater, 2005] Sanchez-Vives, M. V. and Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, 6(4):332–339.

[Sanchez-Vives et al., 2010] Sanchez-Vives, M. V., Spanlang, B., Frisoli, A., Bergamasco, M., and Slater, M. (2010). Virtual hand illusion induced by visuomotor correlations. *PloS one*, 5(4):e10381.

[Schmalstieg and Schaufler, 1999] Schmalstieg, D. and Schaufler, G. (1999). Sewing worlds together with seams: A mechanism to construct complex virtual environments. *Presence: Teleoperators and Virtual Environments*, 8(4):449–461.

[Sciutti et al., 2012] Sciutti, A., Demougeot, L., Berret, B., Toma, S., Sandini, G., Papaxanthis, C., and Pozzo, T. (2012). Visual gravity influences arm movement planning. *Journal of neurophysiology*, 107(12):3433–3445.

[Shadmehr et al., 2010] Shadmehr, R., Smith, M. A., and Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual review of neuroscience*, 33:89–108.

[Slater, 2003] Slater, M. (2003). A note on presence terminology. *Presence connect*, 3(3):1–5.

[Slater, 2009] Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3549–3557.

[Slater et al., 2008] Slater, M., Perez-Marcos, D., Ehrsson, H. H., and Sanchez-Vives, M. V. (2008). Towards a digital body: the virtual arm illusion. *Frontiers in human neuroscience*, 2.

[Slater et al., 2010a] Slater, M., Spanlang, B., and Corominas, D. (2010a). Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. *ACM Transactions on Graphics (TOG)*, 29(4):92.

[Slater et al., 2010b] Slater, M., Spanlang, B., Sanchez-Vives, M. V., and Blanke, O. (2010b). First person experience of body transfer in virtual reality. *PLoS ONE*, 5(5).

[Slater and Usoh, 1993] Slater, M. and Usoh, M. (1993). Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators & Virtual Environments*, 2(3):221–233.

[Spillmann et al., 2013] Spillmann, J., Tuchschmid, S., and Harders, M. (2013). Adaptive space warping to enhance passive haptics in an arthroscopy surgical simulator. *Visualization and Computer Graphics, IEEE Transactions on*, 19(4):626–633.

[Steinicke et al., 2010] Steinicke, F., Bruder, G., Jerald, J., Frenz, H., and Lappe, M. (2010). Estimation of detection thresholds for redirected walking techniques. *IEEE Transactions on Visualization and Computer Graphics*, 16(1):17–27.

[Stoakley et al., 1995] Stoakley, R., Conway, M. J., and Pausch, R. (1995). Virtual reality on a wim: interactive worlds in miniature. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 265–272. ACM Press/Addison-Wesley Publishing Co.

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

[Synofzik et al., 2008a] Synofzik, M., Vosgerau, G., and Newen, A. (2008a). Beyond the comparator model: a multifactorial two-step account of agency. *Consciousness and cognition*, 17(1):219–239.

[Synofzik et al., 2008b] Synofzik, M., Vosgerau, G., and Newen, A. (2008b). I move, therefore i am: a new theoretical framework to investigate agency and ownership. *Consciousness and cognition*, 17(2):411–424.

[Taylor, 2002] Taylor, L. (2002). *Video games: Perspective, point-of-view, and immersion.* PhD thesis, University of Florida.

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

[Tsakiris et al., 2010] Tsakiris, M., Longo, M. R., and Haggard, P. (2010). Having a body versus moving your body: Neural signatures of agency and body-ownership. *Neuropsychologia*, 48(9).

[Tsakiris et al., 2011] Tsakiris, M., Tajadura-Jiménez, A., and Costantini, M. (2011). Just a heartbeat away from one's body: interoceptive sensitivity predicts malleability of body-representations. *Proceedings of the Royal Society of London B: Biological Sciences*, 278(1717):2470–2476.

[Tversky, 1993] Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. In *European Conference on Spatial Information Theory*, pages 14–24. Springer.

[von Helmholtz, 1910] von Helmholtz, H. (1910). *Helmholtz's Treatise on physiological optics*, volume III. Translated from the 3rd German edition and edited by J. P. C. Southall. Dover. Courier Corporation, New York.

[Vosgerau and Newen, 2007] Vosgerau, G. and Newen, A. (2007). Thoughts, motor actions, and the self. *Mind & Language*, 22(1):22–43.

[Walsh et al., 2011] Walsh, L. D., Moseley, G. L., Taylor, J. L., and Gandevia, S. C. (2011). Proprioceptive signals contribute to the sense of body ownership. *The Journal of physiology*, 589(12):3009–3021.

[Wang et al., 2011] Wang, G., McGuffin, M. J., Bérard, F., and Cooperstock, J. R. (2011). Pop-up depth views for improving 3d target acquisition. In *Proceedings of Graphics Interface 2011*, pages 41–48. Canadian Human-Computer Communications Society.

[Wolpert and Flanagan, 2001] Wolpert, D. M. and Flanagan, J. R. (2001). Motor prediction. *Current biology*, 11(18):R729–R732.

[Wolpert et al., 1995] Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232):1880.

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

List of acronyms

1PP: First Person Perspective

3PP: Third Person Perspective

CDR: Control Display Ratio

EEG: Electroencephalography

FBOI: Full Body Ownership Illusion

FOV: Field of View

GSR: Galvanic Skin Response

HMD: Head Mounted Display

IU: Interval of Uncertainty

JND: Just Noticeable Difference

MT: Mean Time

PSE: Point of Subjective Equality

RHI: Rubber Hand Illusion

SOE: Sense of Embodiment

VBI: Virtual Body Illusion

VE: Virtual Environment

VHI: Virtual Hand Illusion

VMT: Visuo Motor Tactile

VR: Virtual Reality

Henrique Galvan Debarba

Date of Birth: Dec 14, 1986
Contact: henrique.galvandebarba [at] epfl.ch
Web page: henriquedebarba.com

RESEARCH INTERESTS

Human-Computer Interaction and Interfaces
Virtual and Mixed Reality
Computer Graphics

EDUCATION

2012 - present : **PhD.** in *Robotics, Control and Intelligent Systems* in progress
École Polytechnique Fédérale de Lausanne | **EPFL**, Switzerland.
Advisers : *Dr. Ronan Boulic* and *Dr. Bruno Herbelin*

2010 - 2012 : **MSc.** in *Computer Science*, with scholarship provided by CAPES
Federal University of Rio Grande do Sul | **UFRGS**, Brazil.
Advisers: *Dr. Luciana Nedel* and *Dr. Anderson Maciel*.

2006 - 2009 : **BSc.** in *Digital Technologies*
University of Caxias do Sul | **UCS**, Brazil.
Adviser: *Dr. Anderson Maciel*.

RESEARCH EXPERIENCE

Sept 2012 – present : doctoral assistant at **EPFL**
Immersive Interaction Group (IIG) and *Laboratory of Cognitive Neuroscience* (LNCO).
Advisers : *Dr. Ronan Boulic* and *Dr. Bruno Herbelin*

Mar 2010 – Aug 2012 : Research oriented master studies at **UFRGS**
Computer Graphics, Image Processing and Interaction Group.
Advisers: *Dr. Luciana Nedel* and *Dr. Anderson Maciel*.

May 2011 : Research Internship at *Université Catholique de Louvain* | **UCL**
OpenInterface project.
Advisers: *Dr. Benoit Macq* and *Dr. Jean-Yves Lionel Lawson*.

Mar 2009 - Dec 2009 : Undergraduate research internship at **UCS**
Voluteer at the Digital Technologies laboratory.
Adviser: *Dr. Anderson Maciel*.

Mar 2006 - Dec 2009 : Undergraduate research internship at **UCS**
New Technologies in the Visual Arts (NTAV) laboratory. Scholarship provided by **CNPq**.
Adviser: *Dr. Diana Domingues*.

AWARDS AND GRANTS RECEIVED

IEEE 3D User Interaction 2016 **Contest winner**

Medicine Meets Virtual Reality 2012 **Poster Prize**

Grant from the Brazilian **Ministry of Culture** with the group LART (University of Brasilia) for a cultural exchange during the **Oncena Bienal de La Habana** (2012)

LANGUAGES

Portuguese (native) - **English** (proficient) - **French** (intermediate)

COMPUTER SKILLS

Programming: R, C/C++, C#, MatLab, Objective-C, Javascript, and Python

GPU Programming: CUDA, Cg/HLSL, GLSL.

Graphical Libs and Game Engines: Unity3D, OpenGL, Irrlicht, Ogre3D.

Mobile Devices Development: iOS, Android (Java SDK and C/C++ NDK)

I/O devices, sensors and displays: Phasespace, Wiimote, PSMove, Kinect, Phantom Omni, touch screens, accelerometer, gyroscope, magnetometer, tracking systems, tiled displays, HMD, CAVE.

PUBLICATIONS

Debarba, H. G., Perrin, S., Herbelin, B., & Boulic, R. (2015, November). **Embodied interaction using non-planar projections in immersive virtual reality**. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology* (pp. 125-128). ACM. (**18%** acceptance rate)

Debarba, H. G., Molla, E., Herbelin, B., & Boulic, R. (2015, March). **Characterizing embodied interaction in First and Third Person Perspective viewpoints**. In *3D User Interfaces (3DUI), 2015 IEEE Symposium on* (pp. 67-72). IEEE. (**32%** acceptance rate)

Grandi, J., Maciel, A., **Debarba, H.**, & Zanchet, D. (2014, May). **Spatially aware mobile interface for 3D visualization and interactive surgery planning**. In *Serious Games and Applications for Health (SeGAH), 2014 IEEE International Conference on* (pp. 1-8). IEEE

Debarba, H. G., Grandi, J. G., Maciel, A., Nedel, L., & Boulic, R. (2013). **Disambiguation Canvas: a precise selection technique for virtual environments**. In *Human-Computer Interaction–INTERACT 2013* (pp. 388-405). Springer Berlin Heidelberg. (**31%** acceptance rate)

Debarba, H., Nedel, L., & Maciel, A. (2012, March). **Lop-cursor: Fast and precise interaction with tiled displays using one hand and levels of precision**. In *3D User Interfaces (3DUI), 2012 IEEE Symposium on* (pp. 125-132). IEEE. (**21%** acceptance rate)

Debarba, H. G., Grandi, J., Maciel, A., & Zanchet, D. (2012). **Anatomic hepatectomy planning through mobile display visualization and interaction**. In MMVR. In *Medicine Meets Virtual Reality 19, MMVR* (pp. 111-115). IOS Press.

Debarba, H. G., Zanchet, D. J., Fracaro, D., Maciel, A., & Kalil, A. N. (2010, August). **Efficient liver surgery planning in 3D based on functional segment classification and volumetric information**. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE* (pp. 4797-4800). IEEE.

Debarba, H. G., Grandi, J. G., Oliveski, A., Domingues, D., Maciel, A., & Nedel, L. P. (2009). **WindWalker: Using Wind as an Orientation Tool in Virtual Environments**. In *Proceedings of the XIth Symposium on Virtual and Augmented Reality, 2009 (SVR)*. (pp. 133-140). Brazilian Computer Society.

EXTENDED ABSTRACTS

Grandi, J., Bernardt, I., **Debarba, H. G.**, Nedel, L., & Maciel, A. (2016, March). **Collaborative 3D Manipulation using Mobile Phones**. In *3D User Interfaces (3DUI), 2016 IEEE Symposium on*. IEEE.

Debarba, H. G., Grandi, J., Maciel, A., Nedel, L., & Boulic, R. (2013, March). **Towards a disambiguation canvas**. In *Virtual Reality (VR)*, 2013 IEEE (pp. 113-114). IEEE.

Debarba, H., Franz, J., Reus, V., Maciel, A., & Nedel, L. (2011). **The cube of doom: A bimanual perceptual user experience**. In *3D User Interfaces (3DUI), 2011 IEEE Symposium on* (pp. 131-132). IEEE.

PARTICIPATION IN ARTISTIC/CULTURAL PRODUCTIONS

OROBORUS BIOCÍBRIDO: Geografismos do Êxtase workshop and exhibition, 2012.
Domingues, D., Miosso, C., Rocha, A., **Debarba, H.**, Donato, C., Oliveira, A..

Presented at the **Oncena Bienal de La Habana**, Habana, Cuba.

Tangible Instants (Átimos Tangíveis) **-** installation, 2012.
Domingues, D., **Debarba, H.**, Miosso, C., Rocha, A., Lucena, T., Marques, M..

Presented at: **Sensorial Experiences** (Vivêncais Sensoriais), Casa da Cultura Percy Vargas de Abreu e Lima, Caxias do Sul, Brazil. Chapter at the "**Vivências Sensoriais-Leituras Múltiplas**" book.

Tagued Idols - installation, 2008.
Domingues, D., **Debarba, H.**, Artecno Group.

Presented at: **8° International Image Festival** (8° Festival Internacional de la Imagen), Manizales, Colombia / **13° National Literary Journey** (13° Jornada Nacional de Literatura), Passo Fundo, Brazil / **7º International Meeting of Art and Technology** (7º Encontro Internacional de Arte e Tecnologia), Brasília, Brazil.

The Cavern of Trance – installation, 2007.
Domingues, D., **Debarba, H.**, Artecno group.

Presented at: **Memories of the Future: Ten Years of Art and Technology in Itaú Cultural** (Memórias do Futuro: Dez Anos de Arte e Tecnologia no Itaú Cultural), São Paulo, Brazil / **A presença do Curso de Artes nos 40 anos da UCS**. UCS CAVE. Caxias do Sul, Brazil.

REVIEWING EXPERIENCE

Conferences : **CHI** (2016) - **VR** (2015) - **3DUI** (2015, 2016) - **CASA** (2012, 2013) - *IHC/CLIHC* (2011) - *CLEI* (2011) - *SVR* (2010 - 2012) - *WRVA* (2010, 2011).

Journals : **IJHCI**

TEACHING EXPERIENCE

Supervision (EPFL)
MSc. dissertation – Samuel Gruner - Yunpeng Zhou
MSc. Project – Fabien Schmitt - Sidney Bovet - Florian Junker
BSc. Project – Sami Perrin - Khoury Jad-Nicolas Elie

Teaching assistant
Virtual Reality EPFL (2014-2016) – C Programming EPFL (2013-2015)
Human-Computer Interaction UFRGS (2011, 2012)

REFERENCES

Dr. Ronan Boulic, Head of the Immersive Interaction Group (IIG), EPFL, Switzerland.
Contact: *ronan.boulic [at] epfl.ch*

Dr. Bruno Herbelin, Center for Neuroprothetics Deputy Director, EPFL, Switzerland.
Contact: *bruno.herbelin [at] epfl.ch*

Dr. Diana Domingues, Founder and Head of Laboratory of Research in Art and TechnoScience (LART), University of Brasília - UnB, Brazil.
Contact: *dgdomingues [at] gmail.com*

Dr. Luciana Porcher Nedel, Associate Professor at UFRGS, Brazil.
Contact: *nedel [at] inf.ufrgs.br*

Dr. Anderson Maciel, Associate Professor at UFRGS, Brazil.
Contact: *amaciel [at] inf.ufrgs.br*