WHOLE-BODY INTERACTION FOR THE ENHANCEMENT OF PRESENCE IN VIRTUAL ENVIRONMENTS

EVENT Lab – Departament de Personalitat, Avaluació i Tractament Psicològic, Universitat de Barcelona

Christoph Groenegress

Director de Tesis: Mel Slater
Tutor: José Gutiérrez Maldonado

Denominació estudis doctorat: Psicología Clínica i de la Salut

A la línia de recerca: Aplicació de Noves Tecnologies
en Psicología Clínica i de la Salut

8 de abril 2010

Acknowledgements

During these years many people helped, influenced and inspired me in a lot of ways and I am grateful to all of them. I wish it was possible to thank all those people here. First of all I would like to thank my supervisor, Mel Slater. Working with him I enjoyed the maximum freedom I can imagine to develop my own thoughts and ideas while always knowing I could rely on his full support and experience, even during the most difficult of times and from the most exotic locations possibly thousands of miles away from Barcelona. I do not have a second supervisor but I would like to thank here Bernhard Spanlang who filled in this role possibly without quite knowing it. His technical and calming advice proved invaluable throughout all projects and towards the completion of this dissertation. This thesis comprises three experiments and a number of people were involved in them in many different roles. Some people even invested their own free time in order to help me and I would like to take the opportunity to thank all of them.

This doctorate was fully funded by the European Union Future and Emerging Technologies Project PRESENCCIA, contract number 27731.

We developed our first experiment at CITA in Copenhagen, where I stayed for two weeks in November 2006. I am grateful to Mette Ramsgard Thomsen for inviting me to her lab and giving me the opportunity to work with the people in her lab. Without Mette this project would have never happened and her artistic background was instrumental in getting it off the ground. The second person I would like to thank is Martin Tamke who designed and produced the interaction space. Chris Christou's guidance on particle systems and passive stereo projections was also vital for realizing this work.

The second experiment was a collaboration between g.tec and the EVENT Lab. I would like to thank Christoph Guger, the head of g.tec, for his invaluable experience and help in completing this project. Clemens Holzner was the person I did most of the – remote – development with and it took a few months until we finally met. It was a pleasure to develop such a big project together with him. When I unexpectedly had to abandon the experiments, Anna Bellido took over and I am very grateful to her for helping me out during this time.

For the third experiment I am thankful to Robert Leeb for his efforts in teaching my how to write effective Simulink code. Andrea Brogni's knowledge on physiology and other things were also most helpful to the completion of the project. Thanks again to Christoph Guger and Clemens Holzner for troubleshooting the g.tec hardware and software with me. Finally, Daniel Pérez Marcos gave me great advice on experimental design. Mar Gonzalez ran most of the experiments and I am very thankful to her for thus allowing me to concentrate on writing this thesis.

Other people who more than deserve to be here are the three other post-docs in the EVENT Lab, namely Iason Kastanis, Jean-Marie Normand and Daniel Pérez Marcos. In particular, I would like to thank Iason and Jean-Marie for reviewing Chapters of this thesis along with Bernhard, and Daniel's experience with running experiments was more than helpful on several occasions. I would also like to thank all the other people from the EVENT Lab I have not mentioned yet and who made working in Mundet such a pleasant experience: Anna Bellido, Ausiàs Pomés, Elias Giannopoulos and Raphael Carbonell. Finally, other people I would like to thank for professional support are Doron Friedman, Yiorgos Chrysanthou, Simon Prince, Torsten Froehlich, Franco Tecchia, David Swapp, Vinoba Vinayagamoorthy, Pan Xueni and Vartkes Goetcherian.

Personally, I have met a lot of people who influenced and challenged my thoughts throughout this PhD and I wish some of them could be here to read this. I am grateful to my family Heidi, Vera, Magnus, Klaus, Marcos and Buddy who kept up with my frustrations and encouraged me in my aspirations. I am also grateful to my friends Christian Luebbert, Guido Ening, Kleanthes Kostas Grohmann, Dirk Florian Schaefer, Jens Januschke, Britta Funck and Lynn, Nils Doehler, Marianna Ramírez and Fabienne Nicholas. I would also like to thank Franz Schuette, Reimo Herfort and Henning Watkinson. Without the help and support of all these people I would not be writing these words today.

Abstract

This thesis is concerned with the development and evaluation of whole-body interfaces and their application in virtual reality. In particular we examine two orthogonal elements of presence, namely place illusion and plausibility. The research presented in this thesis comprises four experiments that examine different types of whole-body interaction – physical, physiological or mental - and assess them in terms of place illusion and plausibility.

For the first experiment our hypothesis was that whole-body movements influencing the behaviour of an abstract environment give rise to plausibility even the environment and effects of user actions were previously unknown. The experiment concerned correlations between whole-body movements and a virtual environment in which participants used a Hula Hoop in order to interact with a particle system displayed consisting of hoop-shaped objects that would individually and collectively respond to the participant’s actions. The immersive environment was displayed on a powerwall. The hypothesis was supported indicating that people can quickly adapt to a new environment and experience plausibility.

The goal of the second experiment was to assess the feasibility of using subjective and mostly unconscious physiological response as a means to modify or enhance certain elements of the virtual environment and thus enhance plausibility. The ultimate goal is to use physiology as an additional tool for storytelling, for example in order to modify or enhance the narrative by increasing the tie between a human and events or other elements of a virtual environment. The underlying assumption was that real and recognizable behavioural constituents, in this case physiological responses of the participant, visualized in a virtual environment should be identifiable by him or her. The second experiment thus also addressed plausibility and we explored how unconscious physiological interactions linked to the behaviour of virtual characters can increase the bond between participant and that character in an environment displaying several virtual characters all exhibiting similar behaviour. The behaviour of one of them was controlled by the participant’s physiology while the remaining ones were automated. Our results suggest that own physiological responses visualized through the behaviour of a virtual character cannot be discerned from a pool of similar but automated stimuli.

The third and fourth experiment both aim to quantify whether brain-computer interfaces used as a universal input device for a virtual environment provide a feasible and stable method for interaction and also if their use affects place illusion, the illusion of being located in a virtual environment. A secondary aim was to demonstrate the feasibility of virtual reality for rapid prototyping of a smart home containing fully automated appliances that can be controlled remotely. In the first of the two experiments we measured performance and subjective presence scores of 12 participants using a P300 brain-computer interface to navigate and interact with objects in a virtual apartment. In the second experiment we collected subjective reports on presence from 12 participants in the same environment, although this time it was controlled via a combination of wand navigation and gaze-based object selection techniques. Our results clearly indicate that when operating the environment via a P300 brain-computer interfaces participants' place illusion are significantly lower, possibly due to high mental workload.

The research presented in this thesis examined physical, physiological and mental means of interaction with virtual environments. The overall conclusion is that while most tools developed yield adequate performance levels, not all of the assessed environments give rise to the desired effect of enhancing presence and this is true in particular when using a brain-computer interface possibly due to high mental workload and split attention.

Resumen

En esta tesis se evalúan nuevas interfaces corporales y su aplicación en realidad virtual. En particular se examinan dos elementos ortogonales de del concepto de "presencia": la "ilusión de lugar", que es la sensación *ahí ("place illusion"*, en inglés) y verosimilitud ("plausibility", en inglés). La tesis se estructura en cuatro estudios examinando varios tipos de interacción – física, fisiológica y mental – que se evalúan en términos de *"place illusion"* y "*plausibility"*.

En el primer estudio la cuestión es si movimientos corporales que influyen un entorno abstracto pueden dar lugar a una sensación de verosimilitud aun cuando ni el entorno ni los efectos de las acciones se conocen a priori. El estudio correlaciona los movimientos físicos de la persona con un entorno virtual en el que los sujetos usan un *hula hoop* para interactuar con un sistema de partículas que consta de objetos de forma de aro que responden a las acciones de manera colectiva e individualmente. El entorno inmersivo es proyectado en una pantalla. Los resultados del estudio indican que uno se puede adaptar fácilmente a nuevos entornos y experimentarlos de manera verosímil.

El objetivo del segundo estudio es la evaluación del uso de las reacciones fisiológicas que son objetivas e inconscientes para modificar varios elementos de un entorno virtual y por lo tanto aumentar la sensación de verosimilitud. El objetivo principal es el uso de señales fisiológicas como herramienta adicional para controlar la narración, por ejemplo para intensificar el vínculo entre la persona y ciertos aspectos del entorno. Se supone que las variables comportamentales de una persona visualizadas en un entorno virtual deben ser identificables por esa misma persona. Por lo tanto, el segundo estudio también trata de verosimilitud y explora cómo las interacciones fisiológicas entre la persona y el avatar virtual pueden amplificar la conexión entre el sujeto y el avatar en un entorno complejo, donde se visualizan varias personas virtuales con comportamientos parecidos. En este caso, un avatar es controlado por la fisiología del sujeto mientras que el comportamiento del

resto de avatares está automatizado. Los resultados indican que las propias señales fisiológicas visualizadas a través de las acciones de un avatar no se pueden distinguir del comportamiento de otros avatares cuyas acciones son automatizadas.

El tercer y cuarto estudio intentan cuantificar si es posible utilizar una interfaz cerebro-ordenador como dispositivo de entrada para controlar un entorno virtual. La cuestión es si este método proporciona resultados estables y también si tiene algún efecto positivo respectivo a ilusión de lugar. Además se pretende demostrar que la realidad virtual es una tecnología adecuada para el desarrollo rápido y eficaz de “casas inteligentes”, equipadas con aparatos automatizados que se pueden controla a distancia. En el primero de los dos estudios se evalúa el rendimiento de 12 sujetos utilizando una interfaz cerebro-ordenador P300 para navegar e interactuar con dispositivos en un apartamento virtual. El segundo estudio se trata del mismo entorno pero en lugar de interactuar con una interfaz cerebro-ordenador los objetos se pueden seleccionar a través de la vista, es decir, dirigiendo la mirada hacia el objeto. En ambos estudios se comparan las respuestas subjetivas de presencia experimentada por los sujetos. Los resultados indican que con la interfaz cerebro-ordenador la sensación de presencia es considerablemente menor comparada con el segundo estudio.

La tesis presenta cuatro estudios sobre tres tipos de interacción en entornos virtuales: físico, fisiológico y mental. La conclusión general es que no todos los tipos tienen el efecto deseado de aumentar la sensación de presencia.

Preface and List of Publications

Two projects presented in this thesis are the result of collaborations with other laboratories, notably the work on brain-computer interfaces presented in Chapter 5. This project was a close collaboration between g.tec OEG and the EVENT Lab and the technical work was carried out with Christoph Guger and Clemens Holzner, with some supervision by Günter Edlinger. All the publications relating to the more technical BCI aspects and user performance were evaluated and analyzed by g.tec, so that the resulting publications were first-authored by a researcher of the company. A study on presence and use of brain-computer interfaces within this project was carried out in June 2009 and the results have been accepted for publication in a special issue of *Presence: Teleoperators and Virtual Environments* on *Brain-Computer Interface (BCI) Systems in Virtual Reality Environments*.

The second project that was the result of collaboration was the one presented in Chapter 3 and the work was developed in the CITA Lab in Copenhagen together with Mette Ramsgard Thomsen and Martin Tamke.

All the publications that have already been published as a result of or during this doctoral research are printed below:

Groenegress, Christoph; Spanlang, Bernhard; Slater, Mel: The Physiological Mirror—a System for Unconscious Control of a Virtual Environment through Physiological Activity: *The Visual Computer* (2010) doi: 10.1007/s00371-010-0471-9.

Groenegress, Christoph and Slater, Mel: Effects of BCI use on Presence in Virtual Reality. *Presence*, 19:1, 1-11 (2010).

Groenegress, Christoph; Thomsen, Mette Ramsgard and Slater, Mel: Correlations between Vocal Input and Visual Response apparently enhance Presence in Virtual Environments. *Cyberpsychology & Behaviour, 12:4, 429-*431 (*2009*).

Groenegress, Christoph; Slater, Mel; Tamke, Martin and Thomsen, Mette Ramsgard: Spinoff - Transferring Energy between Real and Virtual Worlds *Eurographics, 97-100 (*2007).

Edlinger, Günter; Holzner, Clemens; **Groenegress**, Christoph; Guger, Christoph and Slater, Mel: Goal-oriented Control with Brain-Computer Interface. *Proceedings of HCI 2009, 732-740 (*2009).

Guger, Christoph; Holzner, Clemens; **Groenegress**, Christoph; Edlinger, Günter and Slater, Mel: Brain-Computer Interface for Virtual Reality Control *Proceedings of ESANN 2009,* 443-448 *(*2009).

Guger, Christoph; Holzner, Clemens; **Groenegress**, Christoph; Edlinger, Günter and Slater, Mel: Control of a smart home with a brain-computer interface. *Proceedings of the 4th International Brain-Computer Interface Workshop,* 339-342 *(2008).*

Edlinger, Günter; Krausz, G.; **Groenegress**, Christoph; Holzner, Christoph; Guger, Christoph and Slater, Mel: Brain-Computer Interfaces for Virtual Environment Control. *13th International Conference on Biomedical Engineering, 3*66-369 (2008).

List of Figures

List of Tables

Contents

1 Introduction

The vast amount of research and development within the field of immersive virtual reality (VR) and virtual environments (VEs) concentrates on display technologies, tracking, lighting as well as rendering techniques. Compared to the early head-mounted displays that placed people in environments in a state such that they were legally blind, today we have wide field-of-view, light weight, high resolution head-mounted displays at relatively modest cost, and a host of different projection-based systems, such as the CAVE™ [Cruz-Neira et al., 1993]. The goal of all these technologies is to deliver virtual information mostly so that they can accurately reproduce reality and such that the user perceives and accepts the VE as the dominant reality, a phenomenon that is often referred to as presence or telepresence [Minsky, 1980]. Virtual Reality (VR) has potential benefits for the study or treatment of various disorders [Wiederhold and Wiederhold, 2009] and psychotherapy [Riva, 2005]. Other areas of interest include rehabilitation [Rose et al., 2005], training for hazardous environments [Brooks, 1999] or otherwise critical skills such as surgery [McCloy and Stone, 2001]. In this document we will make use of both terms, VR and VE, however we do not regard them as interchangeable. By using VR we will generally refer to a virtual reality system and imply its hardware and setup, the use of VE implies reference to the content and the experience of a VR system.

These examples demonstrate that the industry has developed trust towards this technology and the assumption is that it yields decent results to comparable real-world situations. Some questions that arise in this context however are whether this assumption is valid at all and under what conditions it holds. For example, people may have a different sensation on the vividness and realism of when experiencing the same VE and sometimes and it may be the case that the same person responds different when exposed to the same VE twice. Clearly, the environment does not change but somehow the illusion that what is happening is real might be somehow affected. If it is wrong to assume that VR can deliver virtual environments realistically and that people also respond to them as if they were real, then it cannot be used in the areas described above. The goal of this thesis is to examine some novel interaction techniques centred around the human body that can be exploited in order to enhance some aspects of presence.

In this thesis we aim to contribute to the fields of presence and whole-body interaction by first establishing a connection between the two. Our main focus, then, lies on presence and the application and evaluation of some theories to enhance the sense of presence. We do this by developing a series of distinct whole-body interfaces (WBI) that each deal with a different aspect of a recent approach to presence [Slater, 2009]. Our aim is to not only investigate truly physical but also more abstract bodily actions and these include mental and physiological activity. Thus we consider aspects of whole-body interactions in isolation and evaluate each of them as a general tool for enhancing varying attributes of presence. On the one hand, we will consider the relationship between a WBI and the illusion of being physically located in a VE, while on the other we evaluate how correlations between events in the VE and one's own actions can enhance the realism of the environment.

Most technologically available interaction in virtual environments (VEs) to date, either human-object or human to virtual human – in fact in any computer-augmented environment – involves selection and basic manipulation of (virtual) objects plus the action of locomotion to navigate and position oneself in the dedicated space. Regarding human to (virtual) human interaction, a similar repertoire of actions exists that often exclude the most common aspects of our day-to-day social interaction: speech and facial expressions [Bowman et al., 2004].

Nonetheless, going beyond the Xerox Parc-inspired metaphor of the WIMP interface (Windows, Icons, Menus, Pointing Device) [Shneiderman and Plaisant, 1987] for standard two-dimensional human-machine interaction, in three dimensions the VR interface used in the late 1980s and early 1990s, the six-degree-of-freedom Wand, remains the typical means for effecting interaction with virtual objects in the twenty-first century. This leaves the human body tremendously undervalued in terms of its actual capabilities even for simple operations. This problem is not a new one: van Dam, for instance, observed that advances in interface design are not progressive and continuous like other computer-related disciplines. They rather operate in cycles where, after rapid changes, there are long periods of stability, without many changes [van Dam, 1997].

At a point where VEs are becoming ever more complex and richer in affordances, we are arriving at a situation where the common interaction devices are becoming too simple. While head tracking has effectively become a requirement for many VR installations it enables us to naturally perceive and view a VE. However in spite of this being a powerful tool it does not allow us to manipulate objects and in terms of manipulation the human body is enormously underrepresented, reduced to the size of a single handheld device with only a handful of modes of action. Although we agree that a wand usually conveys the most basic and common operations that can currently be carried out in a VE, we argue that its use does not qualitatively contribute to the experience and in the past this has been a motivation to introduce other devices that intend to deliver more natural ways of action. For example, regarding locomotion there are treadmills [Darken et al., 1997], walking in place [Usoh et al., 1999], pads [Bouguila et al., 2004] and movable tiles [Iwata et al., 2005].

In many ways, the wand hinders a person from connecting with a VE in a natural way since it introduces an inherent sense of artificiality and leads to physical and bodily alienation with respect to the virtual. This does not only affect participants' cognitive and behavioural performance but we also argue that it inhibits their willingness to transport to the virtual reality and consequently suppress the stream of sensory input – even if it is consistent. The use of tools in the real world can lead to a change in perception of our surrounding space [Berti and Frassinetti, 2000] and continuous use can even manifest in changes in the brain and the tool's temporary incorporation in the body image [Kitamura et al., 2003, Imamizu et al., 2000], so it is possible that this also holds for all kinds of tools in VR. However, neither real nor virtual tools are known to provide us with an important capability that our real bodies do provide us with, namely agency. Agency is attributed to both awareness of one's body and action [van den Bos and Jeannerod, 2002], and while the latter is provided by tools such as the wand, the former clearly is not. Further evidence for this is presented in [Tsakiris et al., 2006] where passive body movements during a variation of the rubber hand illusion [Botvinick and Cohen, 1998] produce a fragmented perception of our own body, in contrast with active movements which induce a richer sense of body ownership

Interestingly, interaction devices for computer games are often modelled from real artefacts in order to improve game play and enhance the general feel or the sense of reality of the game; there actually seems to be a long tradition in attempting to compensate for such problems: The steering wheel or the Sony Playstation™ compatible dance pad[1] are examples of this. More recently, in 2007 and 2008, Nintendo's Wii™ console[2], for instance, shows in a compelling way that it is possible to employ handheld devices for interaction that more closely resemble natural action as performed in the real world and subsequently they have produced such a product for a mass market and the goal is to make the experience as natural and realistic as possible. See Figure 1-1 for some examples.

Figure 1-1. Four examples of how natural and realistic playing with the Wii is. Top-left: person playing tennis on a Wii. Top-right: boxing game. Bottom-left: bowling. Bottom-right: dancing exercise using the WiiFit.

[1] http://www.us.playstation.com/PS2/Games/Pump_It_Up_Exceed
[2] http://wii.com/

When one looks at simulators such as vehicle or flight simulators, the use of real objects as user interfaces (UIs) for VR is even more ubiquitous and as we can see in Figure 1-2 and Figure 1-3, the resemblance between real and "virtual" is more than striking. The goal of these technologies is to ensure a high skill transfer from the controlled VR and reality.

Figure 1-2. Image of a Boeing 747-400 Flight Simulator Cockpit.

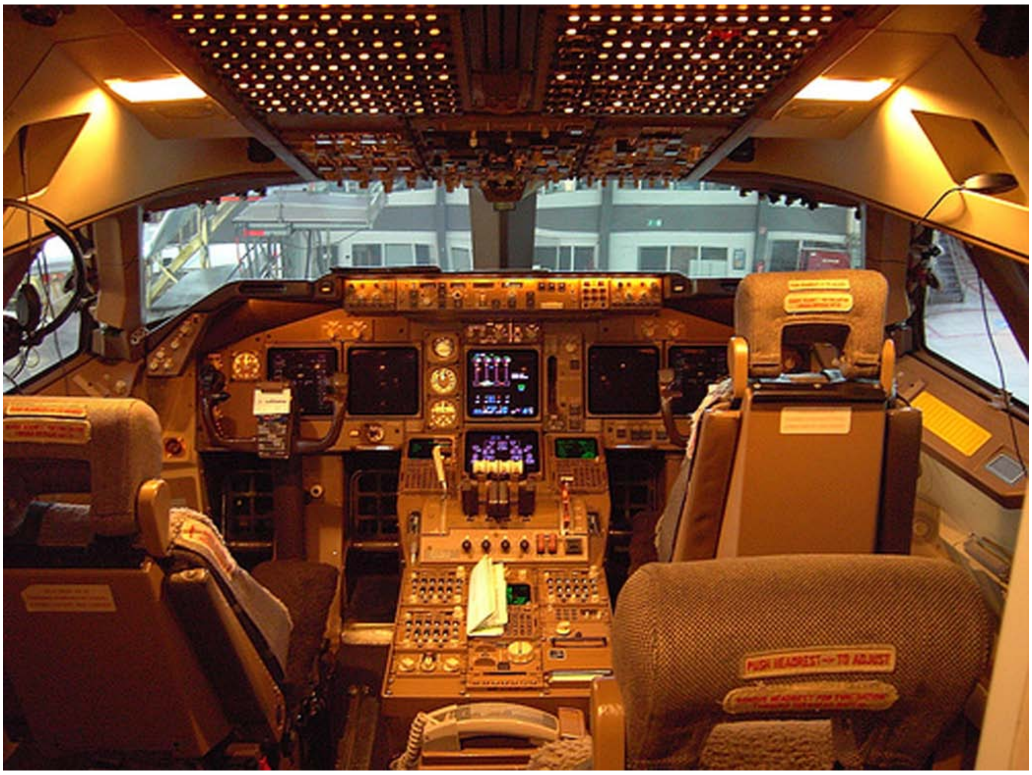

Figure 1-3. Image of a Boeing 747-400 Cockpit.

Although such environments have become quite sophisticated beyond the UIs and the ergonomics they entail, traditionally there is little or no connection with the VR research community [Brooks, 1999].

One reason why developers are keen to re-build reality (in terms of UIs) in VR is because they know they work and also because there is great confidence that such tools increase the perceived realism of the experience, whatever that may be. Presence is the study concerned with the human response to exposure to VR and other media and it has been nearly three decades since Minsky defined the concept of telepresence, from which presence is essentially derived [Minsky, 1980]; we will review the concept of presence in Chapter 2. Presence relates to all the above because it is one tool allowing us to verify the quality and the value of a VR technology or an interaction paradigm. If there was no evidence that people respond to mediated environments as if they were real then it would make little sense to invest in their research and development.

1.1 Research Problem

One of the main problems of presence is that no theory exists that accounts for all observed phenomena or all types of technology used. It has partly been muddled by excessive usage where mediation is often equated with presence. For example, presence has been used to evaluate e-commerce applications [Bente et al., 2003], transmission of emotions over the internet [Tatai et al., 2003], presence when reading fiction [Gysbers et al., 2004], interactive narrative [Pinchbeck and Stevens, 2005], instant messaging [Hwang and Lombard, 2006], blogging [Blasi et al., 2008], evaluating the portrayal of romantic relationships in film [Black et al., 2008] and to assess performing digital art [Bertoncini et al., 2008].

While these are all potentially valid areas for studying presence, it should be clear from those examples that “presence” has a different quality and to some extent meaning in each of them. We therefore need to formally distinguish between these phenomena. In our view current existing theories of presence are not consistent enough. We will revisit this problem in more detail in

Chapter 2. A new paradigm [Slater, 2009], also outlined in Chapter 2 considers four elements of presence: immersion, place illusion, plausibility and the body.

In terms of interaction we noted above that good paradigms often reproduce real-world techniques and one obvious way to generate easily and universally accessible interaction schemes is to exploit the natural functions of the human body and translate these into a repertoire of tools for manipulating virtual worlds. There have been a few attempts to provide a framework for such an embodied, whole-body or body-centred type of interaction [Dourish, 2001, Bowman et al., 2004, Slater and Usoh, 1994] but there has been little exchange between these areas, although there has been some work considering the relationship between embodiment and presence [Biocca, 1997, Schubert et al., 1999, Schubert et al., 2001]

In addition a comprehensive study analyzing and comparing the various different ways in which a human can communicate with virtual worlds either actively or passively do not exist to date partly because they require a strong technical expertise from many different fields such as physiology, psychology, neurosciences, engineering and computer science.

Thus, a significant challenge is to find and understand new processes in which human capacities can be exploited. A secondary aspect deals with evaluating their benefit for the enhancement of presence and a tertiary characteristic defines the boundaries of such interactions.

1.2 Research Questions

Given the above remarks the approach taken in the research presented in this thesis is the exploration of whole-body approaches to interaction and their evaluation towards and enhancement of presence. This research extends earlier work by investigating the feasibility of novel interaction techniques based on physical, mental or physiological processes and incorporating these into a framework for presence and whole-body interaction. The questions that we address all revolve around the same goal, namely the enhancement of presence through whole-body interaction.

Our work comprises two laboratory-based experiments and a third one that was carried out during the exhibition of an interactive installation. We address the following questions:

> *Question 1: What emerges from the relationship between one's own intentional physical actions and resulting changes in a VE?*

This question is addressed in Chapter 3 where we evaluate dance-like real-world movements for use in communicating with VEs and their impact on presence. Our intention in this study is to show that presence can still arise if there is a correlation between actions and responses that were previously unknown.

> *Question 2: What are the effects of a human's unintentional and subconscious physiological processes if used for interaction with a VE?*

In Chapter 4 we introduce a system that controls the behaviour of a virtual character by means of physiological processes of a human participant. We evaluate the feasibility and benefits of such a system. The goal is to show that unconscious physical behaviour can be used to control or direct attention towards certain events displayed in a VE and that this leads to enhanced presence with respect to the environment, event or collection of objects.

> *Question 3 and 4: How do interaction techniques purely based on mental activity fare as a means of input in VEs and how are they accepted by humans as a function of presence?*

These questions are addressed in two experiments presented in Chapter 5. The first experiment outlined in Section 5.3 deals with the first question and a comparative study presented in Section 5.4 considers the second question. For this study we will make use of a brain-computer interface (BCI) as a means to control events in a VE. BCIs are discussed in more detail in Section 2.6.

1.3 Contributions

This thesis contributes to the study of presence as well as physical and, to some extent, social interaction in VEs. Our main focus clearly lies in experimenting with novel whole-body interaction techniques in order to assess their value for the enhancement of a VR experience and presence. As mentioned above we believe that modality plays an important role in presence but can also lead to a more structured and, ultimately, a richer VR experience. The main contributions of our work are listed below.

First of all, we give a detailed review of the state-of-the-art of VR and presence research. We will also discuss trends in human-computer interaction and compile a framework for whole-body interaction and equate its terminology with that used in presence.
Secondly, in three quantitative studies encompassing novel ways of bodily (i.e. physical, mental and physiological) interaction within Virtual and Mixed Reality (MR) environments, we will evaluate each in terms of presence and related issues as well as performance and feasibility of the device or setup. Interaction is accomplished through either straightforward mapping of functions of the human body but also via unique and novel ways of incorporating aspects or movements of the human body into the interaction scheme. Furthermore, the equipment used throughout our studies ranges from inexpensive, unobtrusive and simple to set up to expensive, obtrusive and complex to set up, so our evaluation of interaction techniques will also consider these aspects regarding present and future use.

The three studies and their empirical questions we are concerned with are:

1. Direct physical interaction in an abstract MR environment to evaluate its benefits for the enhancement of presence. In this study we also attempt to show a direct link between physical action and presence in VEs but also the reverse, namely that if the environment responds

consistently to our own actions it thus contributes to an enhanced perception of the VE.

2. Implicit physiological interaction is a study in which we directly test two hypotheses. First of all we intend to show that it is possible to exploit implicit and largely subconscious physical activities and reflective conditions such as physiological states can be used to make a meaningful connection with a VE and that these could be used as a means to draw people's attention to certain processes in the environment. This study also strongly relates to concepts termed correlational presence and plausibility referred to in Chapter 2.

3. Purely mental interaction using a brain-computer interface (BCI) as a generic input device including locomotion and object manipulation. In this project we demonstrate the feasibility of an all-encompassing device based on mental processes for interaction in VEs. The main contribution of our work is the assessment of its use in VR and its impact on presence.

1.4 Scope

There are many possible ways to experience presence. As we observed in Section 1.1, the idea of presence can in theory be applied to an arbitrary medium and as we will elaborate in Chapter 2, it is difficult to make valid assumptions about a range of media as they often come in different flavours, and observations about one variety may not necessarily be valid in another. It is therefore important to define the range to which our work applies. In this research we deal with immersive virtual environments that primarily addresses the visual sense and, to some extent, the sense of hearing. Clearly, an ideal VR system attends to all human senses simultaneously but there is no such system available for study at present.

In this thesis we essentially develop three complementary methods for bodily interaction and analyse and evaluate them regarding their usability and usefulness within certain application scenarios. The actual goal of each of the projects is not the development of new algorithms in computer science or graphics but rather to test a more recent approach to presence [Slater, 2009]. This is done via experimentation and the purpose of our experiments is to show that whole-body interaction can be achieved at different levels of the body controlling different parameters of the VE.

1.5 Structure

This thesis is structured as follows. In Chapter 2, we will cover relevant background. Chapters 3, 4 and 5 present each of the three studies in greater detail. We discuss design, methodology, apparatus, results and conclusions for each topic separately. Chapter 6 concludes the work by combining the overall results and suggests future work.

In Chapter 2 we will place our research in context discussing work relating to whole-body interfaces and in particular our three approaches. We will cover background current topics in presence research and theory, physical interaction, brain-computer interfaces and physiological feedback. In addition we will discuss the state-of-the-art of presence research in terms of immersion, place illusion and plausibility which we distinguish from historical perspective.

Chapter 3 deals with the experiment on natural whole-body and playful interaction in an unrestricted environment. The work was motivated initially by an invitation to present an interactive installation at an architectural exhibition. It shows how movements using a Hula Hoop can be used to interact with an abstract environment to create a MR environment that is, in principle, enjoyable and physically challenging or at least engaging. The study was evaluated using quantitative measures using questionnaires.

In Chapter 4 we introduce a study on plausibility and correlational presence using physiological feedback. Real-time physiological processes are fed into the VE and affect some aspect of it, in this case the behaviour of a virtual character. Physiological measurements taken were heart rate (HR), galvanic skin response (GSR) and respiration. We summarize ways to extract, filter and analyze each signal in and introduce a system that provides all three in real-time. Our study aims to show that, by visualizing some aspect of their largely unconscious physiological behaviour, humans are still able to associate more with a virtual character or object that closely mirrors the human participant's behaviour compared to one that does not.

Chapter 5 presents two studies in which we attempt to quantify the effects of BCI use on presence. We present a virtual smart home environment that can be entirely controlled using thought alone. The first experiment demonstrates the usability of the system in terms of user performance and shows how combining VR and BCI research can lead to a useful tool for interaction. In the second experiment we examined the effect of using the BCI on presence. In particular we correlate workload and attention required by the BCI with the ability to familiarise oneself with the VE. In order to assess this, we used presence questionnaires in both studies but changed the mode of interaction for the second one. Performance was measured using a task-based approach.

Finally, in Chapter 6, we discuss the implications that are common to all three studies and also points out if and where results contradict each other or can be combined. We also point out potential technological as well as theoretical benefits and achievements giving future directions of work.

2 Background

2.1 Synopsis

In this chapter we will discuss relevant background and state-of-the-art of a diverse range of research fields all relating to, in one way or another, whole-body interaction in VEs. Since this thesis is centred around three distinct projects on human-computer interaction in VEs we will summarize each of them separately before joining all the information to draw a more coherent picture of the material that we will cover.

Our central focus lies on presence research and the application of virtual and mixed reality technology towards more natural and emphatic means of human-computer interaction within such environments. In order to do so, however, it is necessary that we develop a good understanding of some other disciplines such as the neurosciences or psychology.

This chapter is organised as follows. We will briefly summarise and discuss some of the goals of Virtual and Mixed Reality research and the various research strands that have emerged from or strongly contributed to it. Of particular concern to us shall be the areas of computer graphics, engineering, interaction design and artificial intelligence. The section is aimed at two aspects. Firstly, to highlight the state-of-the-art of VR technology: While true VR is not only about seeing, hearing and manipulating things, we will mainly concentrate on those characteristics and omit other research branches on sensory displays such as haptics, smell or taste. While we acknowledge that they should form an integral part of any VR system, a thorough review of these falls outside the scope of this thesis. Also, because the majority of existing VR applications exclusively employ vision, sound and sometimes haptics, it is hard to estimate the effects of other sensory displays. Secondly, while it is true that immersive characteristics such as the quality of the display affect how a VE is perceived – it therefore has a strong measurable effect on presence – it is not equal to presence and thus has to be considered separately [Slater, 2003]. From the point of view of presence research the collection of technological ingredients of a VR system is often called immersive technology and Slater and Wilbur defined immersion as the characteristics of an IVR [Slater and Wilbur, 1997]. Among others, these include underlying technical aspects of the VR system such as screen resolution and size, frame rate of the display, the quality of the rendered image, extent of tracking and latency [Draper et al., 1998].

To set presence and immersion apart, we dedicate a substantial part of the following section to developing an understanding of these two concepts. We will begin by looking at different views on presence since its beginnings in the early 1990s until now. While the leading perspective in the 1990s was rather philosophical, the current theoretical foundation of presence is closely linked to other branches of cognitive sciences, and we will cover necessary material as we go along. We will outline a novel theory of presence which is based on three distinct variables: immersion, place illusion and plausibility [Slater, 2009]. In this vein, we will discuss how other disciplines come into play and we will sketch in more detail some vital characteristics such as experience and perception.

We will then move on to discuss some points relating to HCI, 3D user interaction and the emerging paradigm of whole-body interaction and will also establish some parallels between concepts from these disciplines and presence theory.

Finally, we will highlight the state-of-the-art of areas relating to body-centred interaction, brain-computer interfaces (as input devices for VEs) and VR research where physiological measurements are used and why. Each of the sections is designed to give an overview of the particular field and since there is not much overlap between those areas, they are treated separately.

Body-centred interaction, or whole-body interaction, is a fairly loose term for a collection of approaches to interaction and we will focus on its central themes. We will summarise and compare the main ideas in Section 2.5. Regarding brain-computer interfaces (BCIs) there are two research branches, that both intend to solve the same problems primarily of rehabilitation of physical disabilities but also augmentation of the human sensory system. The state-of-the-art of BCI research relating to VR is presented in Section 2.6 and it is necessary because our study on mental interaction presented in Chapter 5 involves the use of a BCI. Likewise, and with respect to our work on physiology, presented in Chapter 4, as a means to manipulate events, in Section 2.7 we will illustrate examples from physiological feedback and show how our work extends the current paradigm by showing how it fits into the context of plausibility.

2.2 Virtual Reality

Virtual Reality has been defined as a “branch of computer graphics” where an experience is “any in which the user is effectively immersed in a responsive virtual world. This implies user dynamic control of viewpoint” [Brooks, 1999]. In contrast, Ivan Sutherland, the pioneer of VR who invented the first VR display in the 1960s describes “a display connected to a digital computer gives us a chance to gain familiarity with concepts not realizable in the physical world. It is a looking glass into a mathematical wonderland” [Sutherland, 1965]. He also notes that “it should serve as many senses as possible” and that “the ultimate display would, of course, be a room within which the computer can control the existence of matter”. Paraphrased, the first definition therefore says that the fundamental requirements of working virtual reality are that it be responsive, that it has the ability to immerse the user in a virtual world and that the user be tracked such that the world updates correctly responding to movements and actions made by the user. The second one rather states a vision without telling us so much how this is achieved. It merely mentions a display or a room (not exclusively attending to vision) producing a set of computer-generated ideas. While the first and more recent one by Brooks tells us where the roots of VR lie, the latter remarks are much more visionary explicitly stating that it should address as many senses as possible and not only the visual aspect of the human sensory system. Attempting to combine both descriptions we can define VR as follows: VR is an *immersive display technique or technology that attends to a variety and in its ideal realisation all human sensory channels, displaying imagery that is* *consistent across all the displays employed. It also incorporates knowledge about the user's position and movements, some of which can be used and interpreted as commands to the system, in order to adequately present an interactive world that can be of any shape or nature, depicting either real world situations or, indeed, any that is not physically feasible but imaginable*.

Advances in graphics research and display technology are the main factors that have shaped the public perception of VR. As we already mentioned in the introduction of the previous chapter, this concentration of research has led us from highly obtrusive, ineffective display devices with low resolution and field-of-view to high-end, light-weight devices. Also the costs to set up and maintain a VR system have fallen dramatically, so that is has now become possible to build a CAVE-like environment for less than 20% of the cost of a similar system a few years ago. A similar comparison holds for HMDs, although the quality has not so much improved rather than their average size and weight decreased. High-end plasma or LCD displays can be found in many households and even projector technology has entered the mass market and is now widely available. Furthermore, in 2007, growth in global revenues from computer games surpassed revenues generated by the music industry and, by 2008, also outperformed the movie industry[3]. In the film industry, the majority of special effects are based on techniques developed in computer graphics research, which, in this sense, has made the biggest leap among the disciplines contributing to VR in terms of public recognition. Other areas such as haptics and 3D interaction techniques are slowly catching up while yet others, such as the two main "chemical" senses of olfaction and gustation, are barely even addressed nowadays. Many VR systems and applications have a strong tendency to concentrate efforts on the visual display and only recently scientists have begun to develop systems that couple more than one display and addressing.

---

[3] http://arstechnica.com/gaming/news/2008/01/growth-of-gaming-in-2007-far-outpaces-movies-music.ars

Finally, in spite of all the technological progress, a scientific discussion on why VR works and how come we trick ourselves into accepting obviously synthetic environments as real has been inconclusive so far. While there is an open and ongoing debate on what is termed *presence*, there is little consensus over what are its defining characteristics, neither how it can be measured or induced by use of certain technology. This is the topic of the next section in which we will develop an account of a presence from its beginnings in the early 1990s until now.

2.3 Presence

Delivering Virtual Reality realistically and believably is a complex task. This is not only because currently we have only limited knowledge about which components are necessary, which are sufficient and which are redundant when displaying and interacting with a virtual environment, we also have to cope with a problem that reality does not need to cope with at all (at least most of the time): a second stream of sensory input coming from the real environment. Whereas in reality it is almost always clear that all sensory input makes up the same environment in a consistent way, this does not hold for a virtual environment whose stream of sensory data in many cases has to compete with the one coming from reality simultaneously [Slater and Steed, 2000, Slater, 2002]. A true VR system allows a human to be fooled into believing that the dominant reality is in fact the virtual one and not the real world. The main problem that arises is that the reception of the environment is largely dependent on subjective psychological states and reactions to the same VR experience can vary greatly among different people and even among the same person. Presence is precisely this discipline that aims to understand and define variables, boundaries and degrees of realistic VR experiences.

The definition of presence that we shall adopt throughout this thesis is loosely based on [Sanchez-Vives and Slater, 2005]: presence is a subjective response to external stimuli and events happening in a VE as though they were real. This definition does not necessarily agree with other people’s definition presented below. A virtual reality experience that can produce presence therefore is not simply dependent on the degree and the modality of immersive technology employed – such as surrounding, stereoscopic, head-slaved vision, auditory and haptic displays. The need for multimodal integration and displays, for example, clearly stems from the knowledge that our perception in the real world is also multimodal and stimuli are often interpreted and *integrated* according to a variety of sources [Stein and Meredith, 1993]. In addition it is strongly dependent on the correlations between different sensory modalities. The visual sense may be regarded as the most dominant and hence the most crucial element of our (multisensory) experience. For example, a person experiencing a virtual fire blazing in front of him while really standing in an air-conditioned laboratory will receive wildly contradictory input from his visual (and possibly auditory) senses and his skin receptors for hot and cold. Although there is anecdotal evidence that people, when confronted with such a scenario, do start to feel illusory higher temperatures, there is no scientific evidence that it does so simply because the brain “overrides” contradictory input from other modalities and succumbs to the (dominant) visual modality [Spanlang et al., 2008]. Much more importantly, there is no evidence that this occurs for the majority of people, and if a VR system cannot guarantee to deliver roughly the same experience for the majority of time and people then it becomes unpredictable rendering any scientific results and inferences about its use problematic.

Coming back to the problem of contradictory sensory information coming from real and virtual environment how is the stream to respond to as real and which to ignore determined by participants? Or is it possible to be present in two different spaces at the same time? Are the processes involved in this decision making conscious or unconscious? Are there ways to manipulate and influence this decision? What technologies and what techniques are necessary to induce the feel of realism in a person given a stream of artificial sensory input?

Presence is the study concerned with issues relating to these questions and one notion is that it is the sense of “being there” in a VE instead of the real physical environment in which the user’s body is actually located [Held and Durlach, 1992, Sheridan, 1992, Ellis, 1996, Draper et al., 1998, Slater et al., 1998]. There are numerous other terms in use to describe the sense of presence under many different conditions. Other metaphors, for example, involve: being here [Sheridan, 1992], being there [Heeter, 1992, Sheridan, 1992], willing suspension of disbelief [Steuer, 1992], distrust of naïve belief [Blake et al., 2007], perceptual illusion of non-mediation [Lombard and Ditton, 1997] etc. A major problem in this debate is that contributions come from a wide range of academic fields, most notably computer science, engineering, psychology, philosophy, cognitive science, the arts and more recently also neurosciences. This entails that terminologies that are in use are heavily influenced by academic background and their use in one area may be restricted and cause them to be interpreted ambiguously or incorrectly in others.

Unclear terminology and contradictory opinions in the community may be the two main reasons why there is an ongoing debate over which parameters affect presence, how it can be produced and its occurrence measured [Lessiter et al., 2001, Witmer and Singer, 1998, Slater, 1999, IJsselsteijn et al., 2000]. Some of the literature claims, for example, that subjective measurements and questionnaires are a valid tool [Prothero et al., 1995] while others deem them unstable [Freeman et al., 1999]. There are also claims that questionnaires alone cannot assess presence and that subjective quantitative data are generally overemphasized in VR while other types of data such as physiological are often neglected in favour of them [Slater and Garau, 2007, Slater, 2004]. In either case, care should be taken when designing a presence inventory [Thornson et al., 2009]. Furthermore, questionnaire data does not necessarily give insight into whether presence actually exists as a quantifiable brain activity that relates to the actual experience [Slater, 2004, Usoh et al., 2000].

An alternative approach for measuring presence was proposed in [Slater and Steed, 2000]. A break in presence (BIP) takes place when a participant in a VE becomes aware of the real world and thus stops attending to the virtual stream of information. In today's VR systems these BIPs occur naturally and frequently. Stepping over a cable for example might cause one to become aware of the real-world surroundings. Viewing the edges of a CAVE wall or actually bumping into them are two more examples. This entails that a participant can switch back and forth between two streams of sensory input, either real or virtual. Whether this behaviour is only an unconscious reaction to an external event or whether it can be controlled is currently not known however. Physiological measures have been used successfully in the past to quantify BIPs [Slater et al., 2006b, Brogni et al., 2003, Slater et al., 2003]. Physiological measurements are discussed in Section 2.7 below. BIPs have also been proposed as a standalone measure for usability [Steed et al., 2005].

A first broad and comprehensive account of the different concepts, theories and possible variables influencing presence was attempted by Lombard and Ditton [Lombard and Ditton, 1997]. They first outline how presence can be conceptualized in terms of social richness, realism, transportation (*being here*, *being there*, *being together*), immersion and a medium as a social actor and then move on to assess what factors should be included in a theory of presence. Among others they identify immersive characteristics addressing all of five human senses, degree of interactivity, obtrusiveness and type of medium but also characteristics defining the contents of a VE such as its (social) realism. A third deciding factor is user-dependent and relates to previous experiences and to his *willingness to suspend disbelief*.

Furthermore there are different “types” of presence experiences. Physical or personal presence refers to what we have mainly been discussing in this section. Social presence [Heeter, 1992] and co-presence refer to the degree of awareness of the presence of an interaction partner, either real or artificial [Garau et al., 2005, Pan et al., 2008], and in the former case either co-located [Heldal et al., 2005a, Heldal et al., 2005b] or remote [Schroeder et al., 2001]. This is a concept going back to Short and colleagues [Short et al., 1976] and it has been studied in a variety of scenarios ranging from the design of intelligent social agents [Lee and Nass, 2003, Swartout et al., 2006] to studies on interactive drama [Dow et al., 2007]. Relational presence [Maguire and Connaughton, 2006] is a four-dimensional model relating to social presence. It can be used to grade presence from “fully present” to “absent”. Also, it is not exactly clear whether social presence deals only with the perceptual or the emotional aspects as well [Vanden Abeele et al., 2007]. There have also been few attempts to qualitatively assess the state of presence [Garau et al., 2008, Turner and Turner, 2007]. Some recent studies also employ physiological measurements to assess the sense of presence [Meehan et al., 2002, Slater et al., 2006b, Brogni et al., 2003, Slater et al., 2003].

Correlational presence is a more recent hypothesis stating that humans are more likely to respond to VEs as if they are real if these in turn respond to the human as if they were real [Gillies and Slater, 2005]. This paradigm relates to the definition of social presence by Heeter [Heeter, 1992]:

> "*The premise of social presence is simply that if other people are in the virtual world that is more evidence that the world exists. If they ignore you, you begin to question your own existence."*

Evidence for this has been found in social interaction [Pertaub et al., 2002] as well as in correlations between humans and the environment in general [Groenegress et al., 2009b].

Some scientists argue that higher fidelity of sensory displays (i.e. immersive technology) is sufficient to provoke a higher degree of presence [Biocca and Levy, 1995, Steuer, 1992, Zeltzer, 1992], while more recently it has been suggested that presence is grounded in continuous action and objects in the world are represented in terms of their potential use for action [Zahorik and Jenison, 1998, Flach and Holden, 1998, Schubert et al., 1999]. One step further from that, scientists have called for a more rigorous neuroscientific approach to presence studies [Sanchez-Vives and Slater, 2005, IJsselsteijn, 2002, IJsselsteijn, 2005] suggesting that more input from this field is fundamental for a deeper understanding of all the processes and variables involved in the formation of presence ranging from low-level mental processes to higher-level psychological states and reflections and emotions. IJsselsteijn for instance notes that by attempting to incorporate phenomenology – a study concerned with enabling objective observations to be made from subjective experiences – mental processing and underlying brain mechanisms into presence studies can complement existing work. A phenomenological approach is similar to the being-in-the-world view held by Zahorik and Jenison or Heeter's subjective experience [Heeter, 1992] originating from Husserl or Heidegger's work on existence. Other aspects of a new theory of presence could draw from neuroscientific work on perception of space [Rizzolatti et al., 1997] and the malleability of the body image [Botvinick and Cohen, 1998, Armel and Ramachandran, 2003, Gallagher and Cole, 1995]. The effects of the rubber hand illusion, for example, as first discussed by Botvinick exist in VR as well [Slater et al., 2008], and this could dramatically open up new opportunities regarding the perception of virtual bodies and also haptics.

Neural processes involved when using real-world tools [Imamizu et al., 2000, Kitamura et al., 2003, Maravita and Iriki, 2004] can furthermore give rise to enhanced interaction paradigms for VR and ultimately help to deliver a higher degree of realism. Attention and hypothesis selection are other factors that might contribute to the sense of being there. When interacting with a mediated environment, the brain constantly processes sensory information coming from that mediated environment as well as the real physical environment. Presence therefore also depends on focus and the dominant perception at a given time and according to theory will increase when the immersive environment is perceptually consistent. It should further increase when attentional resources are directed towards the environment facilitating an internal representation of it [Slater and Steed, 2000, Slater, 2002] and as we mentioned above both real and virtual environment may compete in this [Draper et al., 1998].

In summary, presence is a field with input from many disciplines aiming to construct a theory about the subjective psychological state of feeling present in an environment. There are many assumptions and differing views about what are the main factors contributing to this or even whether there is a neural correlate of presence [IJsselsteijn, 2002]. As a consequence of all this, the current terminology is not very well-defined and seems to be expanding rather than converging into a neat theory.

2.4 Place Illusion, Plausibility and the Body

2.4.1 Deconstructing Presence

As we saw in the previous section, the main problem with the current theory of presence so far is that it is not very tangible nor complete and its terminology is applied to many different experiences (e.g. haptics, computer graphics) not necessarily referring to the same thing across those fields. Another problem is that publications are dominated by many small segments of research that, at the same time, intend to make valid predictions across many types of media (e.g. VR, film, literature, television). There exist many “types” of presence, for example, presence, co-presence, social presence, correlational presence and so on. All of these can be dependent on different variables and different measurements are needed to expose them. By introducing all this terminology, the study of presence has become somewhat unscientific.

However, there have been few attempts to unify all of these phenomena into a single and consistent theory. As mentioned above, Lombard and Ditton classified presence according to the then existing literature [Lombard and Ditton, 1997], Sanchez-Vives and Slater [Sanchez-Vives and Slater, 2005] and IJsselsteijn [IJsselsteijn, 2005] made preliminary calls for a presence theory that ties up more with findings from psychology and neuroscience.

In this section we will outline a new approach to presence [Slater, 2009] that aims at resolving these issues. It defines a few terms that can be applied more universally to the field of virtual reality and presence. It encompasses terminology from perceptual psychology and consciousness, such as sensorimotor contingencies and qualia. These will be explained in due course, for a more in-depth account of sensorimotor contingencies refer to [O’Regan and Noë, 2001, O’Regan et al., 2001, Noë, 2005]. In this theory presence is essentially composed of four factors. These are immersion, place illusion, plausibility and the body, all of which we will scrutinize in the following sub-sections beginning with immersion.

2.4.2 Sensorimotor Contingencies and Immersion

In Section 2.1 we introduced the concept of immersion. We defined immersive technology as any type of display technology plus a tracking system jointly enabling the experience of computer-mediated environments. By display we mean any type of technology addressing one of the human senses (e.g. visual, auditory or haptic). A digital environment is maintained by a computer and rendered on each display and images for each are generated according to the tracking system, e.g. a 6 degrees of freedom (6DoF) head tracker, which ultimately picks up the motions and actions generated by the human user of the system.

Several ways of delivering an IVR exist today and they typically deliver stereo vision, sometimes coupled with audio and in rare cases also haptic feedback. Two common examples are the head-mounted display (HMD) and another one is the CAVE [Cruz-Neira et al., 1993]. However, both are quite different in terms of the display. The former has one or two displays mounted very close to the eyes and the latter is essentially a large room where typically three walls and one floor are projection screens. A head tracking system captures orientation and position of the user's head which in turn is used to determine and align images so that they are rendered correctly: in the case of an HMD two images are rendered, one for the left and one for the right eye, while in a 4-walled CAVE the image is rendered eight times, once for each eye and screen. Stereoscopic glasses then allow for three-dimensional perception of the scenery displayed on the walls.

Now, previously immersion was defined as the characteristics of the IVR [Slater and Wilbur, 1997] (cf. Section 2.1). Another way to describe an immersive system could be through the sensorimotor contingencies (SCs) that they support. SCs are systematic co-occurrences of body movements with correlating sensory stimuli. They were introduced by O'Regan and Noë [O'Regan et al., 2001] and relate to sensorimotor interactions which are interactions between a human and the environment involving the human's use of motor effectors thereby receiving sensory input. Motor effectors are essentially body movements, head movements, for example, result in changes in gaze direction. A sensorimotor skill is defined as the purposeful motor interaction with an environment in order to willingly modify it. An example would be grasping a cup of coffee in order to drink it. It involves three distinct stages planned motor commands and resulting sensory variations. A sensorimotor skill needs to learned by a human and this type of learning is tightly linked to knowledge and experiences relating to the own body. Consider again the process of picking up a cup of coffee, one might feel the sensation that the cup is warm or hot and that it is a solid as opposed to a soft object. The three stages involved in this process are as follows. The first stage relates to the choice of motor commands that allow one to carry out this action. The second and third stage are the perceived sensory variations (haptic, tactile etc.) as a response to one’s actions and the analysis of the sensory variations. According to this definition, certain objects thus convey intrinsic and objective properties. In terms of visual perception, we have knowledge about how postural and positional changes can affect our perception of an object or a collection of cluttered objects in space. We know how to change our posture and head orientation in order to look underneath or behind an object, for instance. An SC that is supported by an IVE is a valid action and the collection of supported SCs is the set of Valid Actions (VAs).

Valid Actions come in two flavours; they can be those types of actions that result in perceptual changes to the imagery displayed. If head tracking is enabled then turning one’s head results in rotation of the scene, moving in one direction results in translation of the scene and so on. These types of actions are largely analogous to actions in a physical environment that relate to physical movements and perceptual changes in the environment. We define them as the set of Valid Sensorimotor Actions (VSAs). The second type of Valid Action is called a Valid Effectual Action (VEA) and VEAs are those that effect changes in the environment. Examples include grasping and manipulating objects. The union of Valid Sensorimotor Actions and Valid Effectual Actions form the Valid Actions of a VE - the repertoire of actions that consistently effect perceptual or environmental changes.

One result of this classification is that we can now define a taxonomy of immersive and non-immersive systems in terms of their SCs. In an ideal immersive system we can fully simulate systems that support fewer SCs. For example, in a CAVE it is possible to simulate to some degree a desktop environment – vice versa it is not possible because a CAVE is a superset of a desktop computer in terms of supported SCs. A higher level of immersion is thus defined in terms of the SCs supported by the system and a lower-level system forms a proper subset of a higher-level one.

Immersion is completely determined by the properties of the system and displays and interactive capabilities are united in this framework. The effects of using different systems are called *place illusion* and *plausibility* and they are discussed in the following two sections.

2.4.3 Place Illusion

A high-level immersive system supporting an adequate number of SCs similar to a physical environment may result in the illusion of being (physically) located inside the displayed virtual environment. As discussed in Section 2.3, this well-known phenomenon has been defined as presence or telepresence. As it is a conscious subjective experience albeit with repeatable and recognizable characteristics it is hard to measure presence. Behavioural, subjective and physiological measures were already discussed in Section 2.3 and as previously argued in [Slater, 2009], the term presence has been associated with many other meanings and in order to explicitly distinguish the "illusion of being there" from other associations the term Place Illusion (PI) is favoured over presence. PI, like immersion, is a function of the supported SCs but the fundamental difference between them is that PI is strongly dependent on the exploitation of existing SCs whereas immersion is not – the former can vary during an experience while the latter is merely a system description. Extensive use of the SCs supported by a system can give rise to a higher frequency of breaks in PI while simply observing a scene does not. Although this does not imply that breaks in PI do not occur during a more passive use of a VE, the probability that they take place is significantly lower.

In terms of immersion SCs are merely an objective description of the boundaries within which PI can occur and it can be used to distinguish between different-order systems – PI can give rise to the illusion of being located elsewhere. Moreover, in terms of PI it only makes sense to compare experiences from the same class of immersion. Comparing a higher-order system with a lower-order one yields no interesting results as the boundaries in the lower-order one can be probed much more easily and frequently than in the higher-order one and both are qualitatively different experiences.

2.4.4 Plausibility

As opposed to PI, Plausibility (Psi) is “the illusion that what is apparently happening is really happening even though you know for sure that it is not” [Slater, 2009], so it relates directly to events happening in the environment rather than the displayed environment itself and its relationship with the SCs afforded by the system. The main assumption is that if events in the VE directly refer to you than this might increase your bond with it. Such events can be triggered without one’s direct intervention but they relate back to and confirm your own participation in the ongoing events. Such events often relate to social events [Pan and Slater, 2007, Pertaub et al., 2002, Heeter, 1992] but they can also be behavioural including shadows and reflections [Slater et al., 2009]. Note however that a visitor of a VE who is not visible to the environment and other visitors, while being fully able to experience PI through his movements and other actions, is unable to experience Psi because the environment does not know that he is there and thus cannot react to his presence.

In many ways, Psi is very similar to PI but often Psi relates to lower-level mental or physiological feelings or responses. While being consciously aware that what is happening is not real people often cannot help but respond as it was real. This is also where Psi and PI differ: one can have a compelling illusion of being in the place depicted by the IVE (i.e. PI) but without further correlations between the environment and oneself the presented scenario may not appear real.

2.4.5 The Body

In environments such as the CAVE people can see their physical bodies and the idea is that they are transported into the VE with their own bodies which are visibly located there. This characteristic is hardly noticed by many users because they can see their own body and almost act naturally within the VE. The problem is that usually the real body is not tracked apart from head and perhaps hand or arm tracking and therefore it does not have a function other than limited locomotion and head movements. If however a virtual body whose movements somehow relate back to the person was rendered in front of them, this perception might change because all of a sudden the real body receives a virtual presentation, e.g. shadows. Similarly, in the case of an HMD, which completely replaces reality including our own body unless a virtual body is rendered in a similar location to where we would expect our real body to be, we would be left without a virtual representation of our own body. Now, if we have control over this body, if its limbs moves in synchrony to our own limb movements this can lead to a powerful sensation of ownership – a correlation between proprioception and visual exteroception. Since this sequence of events that relate to you it can be argued that, to some degree, it becomes your body and this is also where PI and Psi are linked: your body *is* in the place you perceive yourself to be in, even if this body is only a virtual representation of your real one [Slater et al., 2009]. It is PI because you have the sensation of actually being the depicted space. It is also Psi because your real body movements are represented and reflected upon by your virtual body.

2.4.6 Summary

Presence had previously been defined as the illusion of being there but research suggests that it must be classified more carefully. We have described a new approach to presence dividing it into four aspects consisting of sensorimotor contingencies (SCs), Place Illusion (PI), Plausibility (Psi) and the body. SCs define the characteristics of an IVE. PI relates to the illusion of being located in the VE and Psi to the illusion that the things that are presented are apparently actually happening. PI and Psi are fused in the body in the sense that a degree of ownership over a virtual body and correlations between real and virtual body are necessary in order to perceive either.

2.5 Whole-Body Interaction and Presence

2.5.1 Interaction with digital environments

A crucial component of a good VR system is its capability for action and interaction with virtual objects or characters. Interaction devices themselves form part of the immersive technology of a VR system, however. Indeed, most VR input devices have not undergone many changes in the past fifteen years while new developments in visual displays have been more rapid and their quality has steadily improved over time.

Thus, on one side there is immersive displays, on the other side we have interaction devices that allow us to act and manipulate objects effectively. These two are dependent on the type of environment (e.g. mixed, virtual) we are presenting to a participant which in turn leads to the (desired) human response towards this system as a whole. We therefore believe that the more complex a virtual environment becomes in terms of its immersive capabilities, i.e. the more sensory modalities it addresses coherently and realistically, the more realistic and complete should be the interface and the better it should resemble the ways we interact in the real world.

Some areas of interest regarding user interface design are ubiquitous computing [Weiser, 1991, Weiser, 1993] in which computer processes are embedded in other objects or everyday devices. A number of technologies emerged in the 1990s demonstrating the feasibility of this notion with an interactive display system called Liveboard [Elrod et al., 1992] or a desk that allows for the manipulation of real and digital documents [Wellner, 1993], including physical objects to manipulate digital processes [Fitzmaurice et al., 1995]. The work culminated in the emergence of tangible bits at the MIT Media Lab [Ishii and Ullmer, 1997, Underkoffler and Ishii, 1999, Carvey et al., 2006, Ishii et al., 1999] and since then has been incorporated in new paradigm called embodied interaction [Dourish, 2001] stressing the *participative* property of objects, people and concepts, thus giving them contextual meaning and value. In a similar vein, multimodal interfaces combine different modalities (e.g. speech, touch, vision) in order to increase usability [Oviatt, 2002].

In VR there are efforts to merge real and virtual spaces [Lok, 2004, Whitton et al., 2005] so that human-virtual object interaction becomes more natural. This normally involves some form of ideally unobtrusive tracking of the human body and interpretation of body postures and a repertoire of gestures [Schlattmann et al., 2009, Schlattmann et al., 2007]. Tracking is often based on vision by using marker-based motion capture [Lympourides et al., 2009] or markerless multiple-view cameras and a novel view is generated from set of existing views [Avidan and Shashua, 1998]. In the ideal case the real body shape and motion are reconstructed [de Aguiar et al., 2005]. This reconstructed human can then be registered with the virtual space [Ahmed et al., 2005] so that interaction with virtual objects – though lacking haptics – becomes seamless [Theobalt et al., 2008, Weinland et al., 2006].

Other approaches to whole-body interaction include the fusion of different devices such as gaming peripherals [Sivak and McKinley, 2009]. The idea is to use a number of different peripherals that combined aim to achieve a similar task to reconstruction of body motion and posture.

Regarding the evaluation of interaction devices, a number of requirements and evaluation techniques have been discussed in the literature [Slater et al., 1998, Poupyrev et al., 1997, Bowman et al., 2001, Foley et al., 1995] and there is a general consensus that the following factors play a vital role in determining the effectiveness of a device: Completion time, accuracy and error rate as well as ease of use, ease of learning and sense of presence. While the first three factors are purely quantitative and task-oriented, the remaining three convey a more human-centred, qualitative and subjective side of the interaction. The first three measures can be taken objectively, while to date there are no objective measures for either determining the ease of use, the ease of learning or the sense of presence. Instead, these are usually assessed using subjective measures such as questionnaires or interviews.

Finally, quantitative measures usually require us to compare the numbers with existing data from previously conducted experiments, but they can also be misleading, since they, more than the qualitative measures perhaps, are often more dependent on the cognitive abilities of the user of a device rather than the device itself.

As stated earlier, one goal of VR is to facilitate the creation of realistic and life-like experiences in virtual worlds. Immersive characteristics that influence this are the sensorimotor contingencies (SCs) and the Valid Effectual Actions (VEA) [Slater, 2009] the system addresses and supports. It is widely accepted that interactions within VEs can be divided into a set of elementary tasks and the three most common ones are navigation in order to change viewpoint, object selection and object manipulation [Bowman et al., 2004]. Navigation refers to the simple but essential ability to iteratively change one's respective position or viewpoint within an environment. Elementary navigation is a requirement for the accomplishment of the remaining two tasks of object selection and manipulation and selection is necessary for enabling manipulation to the user. Object selection refers to the set of techniques that enable a person to select one or more objects from a larger set. The majority of operations that can be performed in today's VEs can be accomplished using a wand or other equally flexible devices that can be adapted for different settings. As VEs become more complex however, so do the SCs and Valid Effectual Actions that are supported and new ways have to be found to address them. A current paradigm that relates to these issues is called whole-body interaction and it is aimed at supporting more natural VEAs that address the human body and its natural resources. In the following sub-sections we will outline some concepts and related work associated with this topic, stressing the similarities between plausibility and place illusion on the one hand and whole-body interaction on the other.

First of all, however, we will briefly revise some guidelines for UI design in VR. These are important because they allow us to probe the validity of overall

2.5.2 User Interface Design for Virtual Environments

In the previous section we introduced some general concepts and ideas about user interface design, some of which loosely relate to VR. Now, we already mentioned that VR applications tend to be biased towards the enhancement of perceptual aspects, principally vision, and other tasks are often neglected. Bowman and Hodges were among the first point this out [Bowman and Hodges, 1995, Bowman et al., 2004] and they suggested a branch dedicated to the development of 3D user interfaces (3DUIs) should at least partially adhere to existing rules and principles governing traditional user interface (UI) design. One reason for this was to partly compensate for the lack of effort in that area but also to establish some more rigorous design guidelines for interaction in VR. One of the pioneering contributions to the field of design and interaction was Donald Norman and in his book hey lays out guidelines for the *design of everyday things* [Norman, 1990]. Appreciating that humans interact with a large amount of different objects, media and systems every day, he acknowledges that some of them are simple and intuitive to use while others are simply frustrating and prone to errors. The design guidelines presented in his work are derived from observations and studies about how humans perceive, perform and interact in the real world and claim that a good interface should embrace the way we interact in the real world while letting machines perform their tasks transparently. We will briefly discuss the major elements of these principles that are relevant to 3DUI design. They are affordances, mappings, feedback and constraints. As opposed to ubiquitous and embodied computing briefly mentioned in the previous section, some of these loosely relate to Gibson's *ecological approach to visual perception* [Gibson, 1979]. The central claim of this work is that invariant (visual) information about the external world is collected rather than processed, and *afforded* by objects in the environment. By merely seeing objects therefore they provide us with their intrinsic functionalities and these values and meanings can be directly perceived. The main guidelines are summarised in Table 2-1.

Norman further outlined functions such as conventions, visibility, mental models, modes and action cycles although these are not included in the 3DUI guidelines discussed by Bowman and colleagues. It is important however, that

Table 2-1. Norman's guidelines for UI design adapted for VR [Bowman and Hodges, 1995].

Affordances	**Elements of a tool that perceptually display its purpose. Appropriate affordances lead to appropriate use and meaningful actions**
Mappings	**User input should produce a proportional and coherent response to the system**
Feedback	**Follows from good mappings and its purpose is to display information about what precisely an action has triggered**
Constraints	**Limit the number of possible actions of an object so that its functionality remains transparent**

a good 3DUI adheres to most or all of these guidelines. Mappings should be as natural as possible and draw upon physical analogies and cultural standards which lead to immediate comprehension without the need for specific training. Furthermore, the authors argue that often a more constrained UI outperforms a more general one because it is (more) accurate allowing precise input to the system. Objects should in the same way incorporate constraints as do UIs. Constraints can be found in many existing VEs for so-called universal tasks such as navigation e.g., through physical user motion by using real or virtual treadmills [Darken et al., 1997, Slater et al., 1995a], flying or walking [Usoh et al., 1999, Slater et al., 1995b]. User commands, object selection and manipulation are further categories [Bowman et al., 2004].

2.5.3 Whole-Body Interaction and Body Functions

Recent advances in computing power have resulted in increased interest in processing and incorporating real human motions and gestures (cf. Section 2.5.1) such that the VE experience and particularly the interaction become more natural and easy to adapt to. Another aim of this initiative is to reduce the amount of interaction and tracking devices that have to be carried, held or otherwise worn by a person who wants to use a MR or VR. This paradigm of a diminished user interface has been coined Whole-Body Interaction drawing upon many fields. Reality-Based Interaction (RBI) [Jacob et al., 2008] is a paradigm that relates to Whole-Body Interaction and it is intended for designing interaction styles that build on existing everyday knowledge. These are divided into four categories, summarized in Table 2-2 below.

Some of these categories relate directly to Slater's Valid Effectual Actions, i.e. NP and BAS and also SAS, whereas EAS relates more to sensorimotor actions.

Table 2-2. Four categories of interaction styles according to Jacob et al.

Naïve physics (NP)	**Common sense knowledge about the physical world including naïve understanding of physics**
Body Awareness & Skills (BAS)	**Familiarity with own body, proprioception, coordination skills, bimanual interaction**
Environment Awareness & Skills (EAS)	**Contextual information about position and location**
Social Awareness & Skills (SAS)	**support and visualization of verbal and non-verbal communication to enable co-located collaboration**

A very similar classification was introduced to VR as body-centred interaction by Slater and Usoh [Slater et al., 1995b]. Body-centred interaction refers to a more physical and natural interaction paradigm in which the user of a VE can draw from a pool of natural actions in order to navigate and interact with the environment. In a series of studies they examined the benefits of interaction metaphors revolving around this idea. The ultimate goal of these techniques is to maximize the match between sensory data and proprioceptive feedback. According to Synnott [Synnott, 1993] the body fulfils several crucial functions, see Table 2-3 for a summary.

Note how closely Synnott's classification resembles Jacob's in Table 2-2. Synnott's most crucial observation about these functions is that, even though they are clearly present in our everyday lives, they mainly go unnoticed by humans and are largely subject to subconscious processes – a fact that has been widely acknowledged in the neurosciences [Botvinick, 2004].

Table 2-3. Summary of Synnott's classification of the body's crucial functions.

Embodiment	**The physical embodiment of self**
Interaction	**The medium of interaction, through the use of our bodies we interact with and are able to change the world**
Perception	**The anchor of the self in the sensory world: our sensory organs receive data about the external reality which our mind interprets as perceptions of the world**
Communication	**A medium of communication, it allows us to communicate with other humans through the use of sound and gestures.**
Social Representation	**Recognition and identification of others through their bodies.**

In VR however, the conscious self and one's bodily representation in the virtual space are often separate, restricting the flow of consistent sensory data that one would expect given a certain stimulus. Therefore, when experiencing a VE many events give rise to breaks in presence.

In this respect there are several distinct means of engagement that have to be addressed and possibly dealt with separately. The user of a CAVE-like environment for example will "retain" his real body and all functionalities, although not all of them may be addressed by the system (e.g. haptics) at the same time, so his actions may not correspond to valid effectual or sensorimotor actions. Usually such systems allow at least two sensory streams to overlap, i.e. the visual and auditory, although this differs from a Mixed Reality experience. Using a more exclusive and restricted VR display such as the HMD, which only permits visual (and often auditory) stimuli coming from the virtual environment to be picked up by the human and the real world stream is excluded. Thus, there is no notion of a real body in the latter environment and it has to be replaced by a virtual body, ideally attending somehow to the user's real body movements and posture. It should be clear that there are at least two types of engaging with a VR and that these greatly affect the availability and fidelity of the five bodily functions introduced above. One type, for example an HMD, exclude and sometimes replaces the real body with a virtual representation while another one attempts to transport it into the VE.

Thus, one solution to partly compensate for the mismatch produced by this problem is to closely match natural physical actions with a close-to-natural response in the virtual domain. Two of the three elementary tasks in VR as discussed in Section 2.5.1 are locomotion and manipulation of objects. During locomotion, for example, we often can overcome the problem of limited physical space when exploring a much larger virtual space via somehow approximating real walking. A number of techniques have been proposed centred around walking in place [Usoh et al., 1999, Bouguila et al., 2004, Slater et al., 1995b] or motion compression techniques aiming to introduce imperceptible visual-proprioceptive disparities, for example via continuous small rotations and reorientations of the environment during real walking, so-called redirected walking [Engel et al., 2008, Razzaque et al., 2002, Peck et al., 2008]. These become useful if real walking is to control virtual walking and if real space is more limited than virtual space.

Many concepts that we covered in this section strongly relate to the notions of immersion and plausibility, and to some extent place illusion, albeit not as strongly. These were introduced in Section 2.4 and attempt to fit the theory of presence into a more coherent framework. Whole-body interaction, body-centred interaction and reality-based interaction are concerned with enhancing real body functions towards more natural perception and action in VR. While in Slater's theory immersion and plausibility essentially unify the various categories of interaction styles discussed in this section in terms of sensorimotor contingencies and sensorimotor actions, this is merely a broader and more elegant classification and the others form a valid subset of it.

Furthermore, Synnott's classification of bodily functions draws, from a sociological perspective, an interesting picture about what real and virtual bodies currently lack in VR. Thus far, research has only partially succeeded in connecting the real body to the virtual space regarding all of the above attributes. Agreement between the sensory input coming from the virtual environment and proprioceptive data is often fragmented and low such that a mismatch between the two arises and the brain cannot draw a coherent image of the situation. As discussed before the problem, therefore, lies in the ambiguity that both streams of sensory input pose when combined. The brain's inability to cope with conflicting signals has been widely acknowledged and examined in the arts and sciences. Body illusions such as the rubber hand illusion [Botvinick, 2004, Ehrsson et al., 2005a] or the ventriloquist effect [Alais and Burr, 2004] and others [Ehrsson et al., 2005b] are examples where a coherent response is formed out of conflicting sensory data.

It has been argued that telepresence is some form of out-of-body experience [Rheingold, 1991]. Regardless of whether one agrees with this claim or not, it reveals some crucial insight about how we perceive and experience mediated realities in relation to ourselves: they are an external space that we enter not with our given physical bodies but by utilizing some other form embodiment with its own set of controls and affordances. Thus, the common interpretation is that the real body and the virtual embodiment are not the same but distinct entities, one which is innate to us and which we “leave” once we enter a VE and where we use a different body or at least different means to negotiate our way through the environment in order to be in that space.

Summarizing this discussion, a major problem that arises is that we are currently unable to mediate an artificial world such that we can either fully transport the real body into the virtual domain or sufficiently project it into the real world and only then could we truly claim to have achieved interaction addressing the whole body.

2.6 Brain-Computer Interfaces

2.6.1 Overview

In this section, we will give a brief overview of the history and major breakthroughs of brain-computer interfaces (BCIs), sometimes also termed brain-machine interfaces (BMIs). Regarding our work on BCIs for control of a virtual apartment presented in Chapter 5 it is essential to be familiar with some terminology and techniques in use, although our review mainly covers BCI work relating to VR. BCIs essentially form a radically different approach to interaction with VRs and computers in general. There is growing popularity in using them and demonstrating new methods for human-machine communication.

Since BCIs have only recently emerged as a paradigm for interaction in VR specifically, we have to treat it separately from other models of interaction and most of the literature that we will cover in this section is intended to give an general overview of the workings of a BCI and in particular one method to elicit and measure certain brain activity, the P300 interface. We will outline the major advances in the field with respect to our BCI system for VR interaction presented in Chapter 5.

A BCI is a device that measures electrical activity in parts of the brain. The conventional motivation for using a BCI, broadly speaking, is to allow humans (and primates) to facilitate and express physical and mental activities of any kind through thought – at present the only means available to this end is the BCI. It bypasses conventional motor output pathways of nerves and muscles and can provide control over some task for severely-impaired as well as healthy humans. A standard technique for measuring brain activity is Electroencephalography (EEG) and like other techniques it can be performed either invasively or non-invasively [Wolpaw et al., 2002]. A non-invasive and less powerful approach uses electrodes attached to the outer surface of the skull while an invasive approach, requiring surgery, implants electrodes directly into or close to the grey matter underneath the skull. Another difference between them is bitrate, i.e. the number of bits that can be captured and processed at a time, which is usually much greater for implanted devices. There are several ways to interpret and exploit brain activity in order to achieve a task, briefly outlined in the next section.

A BCI allows its user to control dedicated processes on a computer via thought and it does not require any motor abilities. Even though motor imagery can be used for control it does not imply that actual body movements are needed in order to achieve a task. It is therefore a potential candidate for use across a wide range of domains and physical conditions of humans and, indeed, the most active branch of research is in rehabilitation [Donchin et al., 2000, Birbaumer et al., 2000]. Operations can be carried out by simply imagining them or via a so-called stimulus-response approach (see below), where the appearance of a known stimulus provokes higher activity in certain regions of the brain. Both approaches have shown good and stable results in a variety of conditions and domains. Much work on BCIs is dedicated to motor rehabilitation and a lot of work is carried out in primates [Carmena et al., 2003, Nicolelis, 2003, Nicolelis, 2001, Lebedev and Nicolelis, 2006], where the ultimate goal is to recover lost motor functions due to limb amputation using a combination of artificial prostheses and a BCI system to control it.

In the next section, we will give a brief history and classification of some major breakthroughs of BCI research and types of BCI. However, since the bulk of this research is aimed at motor rehabilitation, we will put strong emphasis on BCI use in HCI and VR and omit work that is less concerned with our central theme.

2.6.2 History, Motivation and Classification

Brain-computer interfaces (BCIs) have become a field of extensive research since the 1990s. Although much early, ongoing and promising work has been carried out with primates, we shall put emphasis on non-invasive studies that focus on humans. There are several reasons for this, most if not all work with primates is primarily concerned with restoring motor functions, while the work presented in Chapter 5 aims to evaluate the BCI as a human-computer interface in realistic VEs.

The two main requirements for a working BCI are data extraction and discrimination. The ability to extract desired parameters in real-time from single neurons or neuron ensembles, a collection of neighbouring neurons is a vital necessity. The second precondition is the possibility to extract useful information from a small fraction of neuron samples that are representative of some brain activity or response. This implies that such processes need to be understood and known to some extent, and also that they actually reflect on the user's state or intent. One major assumption here is that the brain should be capable of adaptation and change with respect to training and learning and thus plasticity needs to be taken into account in the design of BCIs [Nicolelis, 2001, Nicolelis, 2003].

As mentioned above, we can classify BCI research into invasive and non-invasive. An invasive BCI states that the subject or patient undergoes surgery and electrodes are implanted into their skull either directly into or close to the grey matter. Non-invasive approaches attempt to read brain activity by attaching electrodes to the outer surface of the skull. There are several ways to achieve this – e.g. magnetoencephalography (MEG), positron emission topography (PET) or functional magnetic resonance imaging (fMRI) –

although EEG is the most studied and currently most cost-effective and practical technique. Since MEG, PET and fMRI are very costly and obtrusive devices, we will focus on EEG-based research only, also because options for delivering VR using the former three methods are very limited.

Generally speaking however, invasive methods are more powerful than non-invasive ones as the signal quality is usually much higher and less noisy. This is also because the spatial frequency can be much higher compared to non-invasive systems. Non-invasive BCIs, on the other hand, are often preferred simply because they do not require any surgery and are less cost-intensive. The main drawback is that the resulting signal quality is poor because it is reduced by skull, hair and tissue. Nonetheless, non-invasive BCI can in many cases be regarded as a more practical and user-friendly alternative compared to invasive BCI when applied to humans.

A finer classification of BCIs can be reached by separating applications into biofeedback and stimulus-response [Nijholt et al., 2008]. Biofeedback is a method where a person is trained to think of a specific body movement or activity and the resulting changes in brain activity can be classified by a computer program. Much more interesting for our purposes are stimulus-response measures which, simply put, determine and match a given brain activity to a corresponding (administered) event in quantitative terms. An event-related potential (ERP) is the term for characteristic changes in brain activity after sudden events. The P300 (ERP) is an example of such a stimulus-response driven BCI and it refers to measurable changes in brain activity 300ms after a subject has been presented with a stimulus and changes in brain activity can in turn be correlated with this stimulus [Sutton et al., 1965].

In the following two sections, we will give a brief outline of the workings of one of the common methods, the P300 interface and other techniques to for BCI control. For a more complete overview of the various approaches currently in use refer to [Wolpaw et al., 2002].

2.6.3 The P300 Event-Related Potential

The P300 wave is an ERP that occurs mainly in the parietal cortex roughly 300ms after the presentation of an infrequent but somehow expected stimulus. This behaviour has been exploited in the P300 interface for item selection. Symbols, most commonly alphanumeric, are arranged in a rectangular grid and individual symbols (or entire rows and columns) in the matrix are highlighted several times in random order (see
Figure 2-1). The human concentrates on a single symbol and every time it flashes, an EEG response can thus be measured 300ms after the event occurs. The trick is to focus on a single symbol waiting for it to be triggered even though it is not exactly known at what interval it is triggered or what order relative to the other symbols. The brain's accumulated responses for a target and non-target are shown in Figure 2-2 (next page). The exact symbol the user intended to select can be determined by accumulating these responses over several iterations because a single response is often too weak to yield a correct result. The time it takes to complete a single iteration depends on the number of symbols presented on the computer screen. As a guideline, on a 6-by-6 grid of symbols, each symbol is highlighted during a period of about 125ms, resulting in 4.5 seconds per iteration.

Figure 2-1. Example of a P300 Interface for alphanumeric spelling.

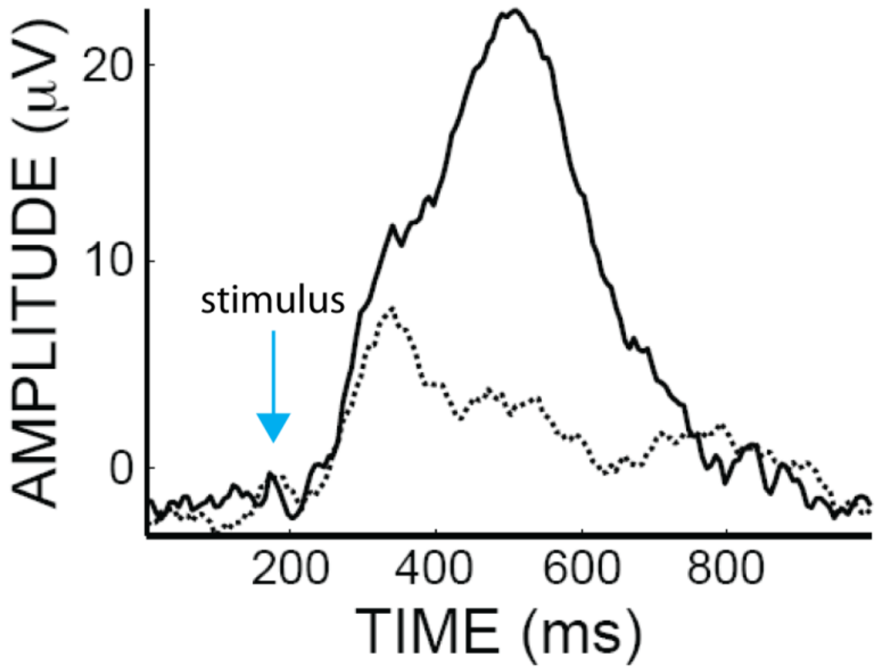

Figure 2-2. Amplitude of brain activity over a 1000ms period. A stimulus is presented at around 200ms. If it presents a target the response will resemble the solid curve, a non-target evokes a flatter response as outlined by the dotted curve.

The number of iterations needed largely depends on experience and attention. For experienced users it can be as low as two. Generally speaking, the larger the number of iterations the safer the method but the more tedious it also becomes. For example, ten iterations require an average of 40 seconds for a single operation (e.g. selecting a letter) to be carried out – much too high a price for healthy computer users who can complete the same task by pressing a button in almost no time at all.

Given that two iterations can yield good results and the amount of time per decision is much lower, a small number of iterations are preferable but not always desirable. A tetraplegic patient who uses the BCI to control items in the household, for example, may choose higher accuracy over speed, because if he is guaranteed to control things himself within a reasonable time frame instead of having to rely on human assistance, this may itself be a rewarding act.

2.6.4 Other types of BCIs

Alternative methods that allow humans to control a device by means of brain activity are briefly presented in this section. These are visual evoked potentials, slow cortical potentials and mu and beta rhythms.

Visual evoked potentials (VEPs) recorded from the scalp have been successfully used to determine a person's direction of gaze [Vidal, 1977, Sutter, 1992, Middendorf et al., 2000]. Their behaviour and application domain are rather similar to the P300 ERP introduced above, however VEPs are produced exclusively in the visual cortex as opposed to the parietal cortex. A collection of visual stimuli is usually presented to the BCI user and these undergo alternations. VEPs arise when the symbol selected by the user is alternating. The VEP method can also be used for spelling.

Slow cortical potentials (SCPs) are slow voltage changes generated in the cortex that occur over 0.5-10 seconds. Positive SCPs are often associated with reduced cortical activation, while negative SCPs often give rise to movements and other functions involving cortical activation. It has been demonstrated that humans can learn to control SCPs [Birbaumer, 1997] allowing them to move an object on a computer screen. This type of BCI is also often referred to as a thought translation device (TTD) [Birbaumer et al., 2000].

Mu rhythms are low-frequency brain activities, around 8-12Hz, in the primary sensory or motor cortical areas during idling activity. They occur in most adults [Pfurtscheller, 1989]. They are associated with slightly higher-frequency activities (18-26Hz) called beta rhythms which are mostly independent of mu rhythms. It has been suggested that these features could be exploited for communication as a decrease in those rhythms is usually related to movement or preparation of movement, a phenomenon that has been labelled event-related desynchronization (ERD) [Pfurtscheller and Lopes da Silva, 1999].

2.6.5 Brain-Computer Interfaces in Virtual Reality and Games

Examples of the use of a BCI as an instrument for human-computer interaction (HCI) have up to date been fairly sparse. The complexity of both the device setup and analysis of brain activity render most activity involving BCIs a difficult and tedious endeavour. In addition there is hardly any work demonstrating its stability beyond the laboratory – a home system offering some functions to one sufferer of amyotrophic lateral sclerosis (ALS) in a fairly stable environment being the notable exception [Vaughan et al., 2006] – though the emergence of a gaming controller claiming to measure brain activity offers a different view[4]. Put simply, the BCI has still not evolved into an easy-to-use tool, that manages to deliver acceptable performance and bitrate while also convincing cosmetically, and much work remains to be done to improve the current situation. Currently it is therefore not an attractive technology for healthy users, although hybrid systems, in conjunction with other devices such as keyboards or mice, are clearly an alternative [Fairclough, 2008].

Brain recordings have been used in a variety of different contexts, for example to monitor a person's performance, attention or fatigue [Huang et al., 2007, Cardillo et al., 2007]. While these examples do not technically provide us with a BCI that "reads" thoughts, they show how the technology is used to pick up on people's EEG data in order to assist them or help augment performance in a number of tasks. Many "real" BCI applications however, in particular those based on the P300 interface, demonstrate that people's ability to control items on a computer screen using thoughts alone [Farwell and Donchin, 1988]. The P300 ERP has been exploited extensively as a spelling device [Guan et al., 2004, Krusienski et al., 2006, Sellers and Kübler, 2006], in which a matrix of alphanumeric letters is presented on a screen and a person can spell words by selecting its letters one by one in the manner described above (see previous section).

[4] e.g. http://emotiv.com

Other applications using the BCI as a mechanism to control computers [Jia et al., 2007, Ma et al., 2007] or specifically VEs [Bayliss, 2003, Bayliss et al., 2004, Leeb et al., 2008] have demonstrated its feasibility in terms of technical requirements. All of them require humans to undergo extensive training periods in order to gain reasonable control over the device and in this context it should be pointed out that BCI control has been identified as a skill that needs to be learned, practiced and maintained [Wolpaw et al., 2002]. Bayliss and colleagues compared the P300 interface in three VR setups: a monitor and a static-camera and interactive scene delivery inside a head-mounted display (HMD). The virtual apartment used for the study offered a total of five actions, and although participants reported better performance in the fully immersive environment, results showed no significant differences between the three display conditions. In games or game-like scenarios, BCIs have been used for binary control in a task involving balancing a virtual character [Lalor et al., 2005] or for control of virtual airplanes [Middendorf et al., 2000].

Another interesting set of experiments was carried out using a method based on motor imagery in order to navigate through a VE. Several experiments showed that imagined movements sufficed to control the trajectory of a virtual character in different environments [Pfurtscheller et al., 2006, Leeb et al., 2004, Leeb et al., 2005, Leeb et al., 2007b, Friedman et al., 2007, Leeb et al., 2007a]. In these studies, EEG activity was captured from the sensorimotor cortex and, over extended training periods, the system learned to classify the participants' motor imagery patterns of hand or foot movement, which in turn could be used for locomotion. Motor imagery was also exploited in controlling a virtual car [Zhao et al., 2009].

A slightly more unusual example combines motor imagery with the so-called rubber hand illusion [Botvinick and Cohen, 1998]. The work demonstrates that motor imagery used to control movements of a virtual arm apparently attached to one's body leads to the illusion of ownership over that arm even though other multisensory correlations such as tactile stimulation were absent during the experimentation phase [Perez-Marcos et al., 2009].

These more recent examples of BCI applications in VR, although bitrates still remain fairly low, may be slowly uncovering a new method for HCI, one that only requires thought to effect actions. Also, the overwhelming majority of BCI studies carried out in VR involve navigation tasks with disabled participants so they are clearly aimed at rehabilitation where VR is only used as a tool to visualize progress and success. Little work has otherwise been done to specifically test BCI performance in VEs.

While it is true that at present only severely disabled people can seriously benefit from the use of a BCI this is very likely to change in the near future. The advent of commercial BCIs for gaming, as mentioned above, shows that there exists the technical potential as well as public interest to use such devices. Next generation gaming devices are likely to adopt this trend and in the medium term they will be used for more conventional activity and partially replace current UIs.

2.7 Physiological Measurements in Virtual Reality

Physiological measurements deal with the capture and analysis of a human's physiological state. In medicine and clinical use these are often used to monitor a patient's health, while psychophysiology is the branch of psychology concerned with the study of physiological responses to behavioural stimuli. The most common measures are summarized in Table 2-4.

Table 2-4. Common physiological measures.

Parameter	Measurement
Cardiovascular	Heart rate (HR), Electrocardiography (ECG), heart rate variability (HRV)
Skin	Skin conductance response (SCR), galvanic skin response (GSR), temperature
Muscle activity	Electromyography (EMG)
Respiration	Respiratory rate (RR)

Brain activity and eye measurements such as eye movements are often considered to be part of this list. There are numerous publications and books relating to physiology, psychophysiology and biofeedback. The goal of this section is however to examine related work on physiological measurements in VR. To this end we will neither discuss techniques nor technology; the interested reader is referred to [Andreassi, 2000, Dawson et al., 2000, Blanchard and Epstein, 1978].

In VR, physiological measurements have only recently started to play an active role. Regarding presence research, for example, heart rate (HR) measures, skin conductance (SCR), skin temperature and Electromyography (EMG) have been used to assess the assumption that an apparently real environment should trigger physiological responses that are comparable to real world situations. In an attempt to make objective predictions about one's *sense of being there* Meehan and colleagues [Meehan et al., 2002] used HR, SCR and temperature to evaluate physiological responses to exposure to a stressful and a non-stressful situations in the virtual pit room [Usoh et al., 1999]. Results from HR and, to some extent, SCR suggest that these environments can cause a likely physiological response and can thus be linked to presence. Similar results were obtained by using passive haptics – a real ledge placed on the floor to enhance the illusion of standing on a real edge – in the same environment [Insko, 2001]. Varying latency also appears to have an effect on the physiological response to a VE [Meehan et al., 2003].

Another experiment deals with the relationship between BIPs (cf. Section 2.3) and physiological responses and shows that HR decreases during intentionally produced BIPs, but also that stress levels increased during social interaction with avatars [Slater et al., 2006b, Slater et al., 2003] which suggest that BIPs can be predicted from events in the physiological data alone.

HR and SCR were also used in a study investigating the human response to virtual humans in a social VE [Garau et al., 2005]. Although physiology was not used as the central element its results are in line with self-reported presence scores on the environment and the avatars' responsiveness. A similar study [Brogni et al., 2007] uses HR, heart rate variability (HRV), respiration and GSR to evaluate physiological responses to photorealistic and cartoon-like virtual characters indicating primarily that being exposed to a VE induces mental and physical stress and more photorealistic characters further increased the participants' stress levels.

In a virtual reprise of Stanley Milgram's experiments on obedience HR, HRV and SCR showed different levels of stress between experimental groups suggesting that people tend to respond realistically at physiological, subjective and behavioural levels when interacting with virtual characters [Slater et al., 2006a].

Electromyogram (EMG) activity was monitored for analysing people's behaviour when walking and balancing on real and virtual beams [Antley and Slater, 2009]. In the latter condition the virtual beam was placed at floor level of a CAVE-like environment while the virtual floor was a few centimetres below the real floor. Analysis of the EMG activity showed that the visual illusion of walking on a beam in the VR condition was sufficient to provide a strong physiological response and both the real and virtual condition had similar results.

2.8 Conclusions

In this chapter we have given an overview of some areas of research that are of interest to this thesis and also gave a state-of-the art account of presence research. The most important outcome in Section 2.2 to 2.4 is a development of a working definition for VR as well as presence. The third major achievement is the summary of the four concepts of sensorimotor contingencies, plausibility, place illusion and the body as a new approach to presence and clearly set these apart from previous presence theories. We also tied this more recent approach to presence discussed in Section 2.4 to concepts from whole-body interaction and 3D user interface design. These ideas were presented in Section 2.5 and we aim to achieve an important connection between those fields. Since our goal is to make assumptions about presence, whole-body interaction and, to some extent, 3D user interface design, there is a need for a tighter connection between the areas, at least on the level of terminology.

The other two areas that we briefly touched on dealt with brain-computer interfaces (Section 2.6) and physiological measurements for VR and presence research (Section 2.7). Regarding the former we need a basic understanding for our work presented in Chapter 5 and some background on the latter is required for some concepts introduced in Chapter 4.

3 Direct Physical Interaction – The SPIN_off Project

3.1 Overview

The overall goal of our three studies presented in this and the following chapters was to develop and evaluate whole-body approaches to user interaction within virtual environments. One obvious way to interact with one's body is through movements of the parts or the whole body and this can be achieved either with or without the use of props. As discussed in Section 2.2 and 2.3 there is evidence that direct use of the body may positively affect plausibility. The main objective of this project was to show that plausibility, as introduced in Section 2.4.4, can be elicited and strengthened through events in a VE that tightly correlate with whole-body actions. Furthermore, our assumption was that even an abstract VE coupled with expressive but unusual movements can elicit such a response after some time.

For this project we were invited by The Royal Danish Academy of Fine Arts to contribute our work to an exhibition that took place in the space of the attached School of Architecture during the month of November in 2006. In collaboration with the Centre for Information Technology and Architecture (CITA)[5] we designed and implemented a Mixed Reality interaction space in which we intended to visualise energy transfer between the real and the virtual realm and thus between the visitor of the exhibition and some virtual entity. The topic of the exhibition was architecture and space, and thus, coming from a VR background, we chose to design a piece that allowed people to transform the real exhibition space into an interactive MR experience using their whole body to attract the attention of virtual entities that would consequently fill the mixed space. Energy was to be generated through body movements and subsequently picked up by and processed by appropriate sensors in order to compute an audiovisual response that would result in a playful and joyful interplay between man and machine.

[5] http://cita.karch.dk

The system was completed onsite in the two weeks prior to the exhibition inauguration, so time constraints were very high for implementing and testing a suitable interface as well as a virtual environment that fit the artistic context.

The work was presented and published in the proceedings of Eurographics 2007 [Groenegress et al., 2007].

3.2 Objectives

The interactive SPIN_off Installation is about the exploration of a reciprocal system in which the user learns to interact with an environment that in turn responds with varying levels of complexity to his/her actions. Small agents with a minimum behavioural intelligence inhabit the virtual space, looking for changes in the usually deserted digital environment. The user creates these changes by entering the space with the Hula Hoop interface. The registered change in density attracts the agents – motivating the user to change his movement in response – the start of an unfamiliar conversation of the individual within the swarm. A relationship emerges as the engagement between user, enhancing his skills and health, and the environment increases in complexity over time. In some sense the participant transfers energy from the real into virtual space.

SPIN_off was designed as a playful environment where users draw a virtual embodiment around their own physical embodiment through physical actions and thereby transfer their own energy generated through real movements into the VE. Working with low-level intelligent systems visualised as hoops that responded to the level of energy they received, the participant would be able to merge the virtual and real spaces around him using this method.

As an input device we chose a Hula Hoop for several reasons. Being a toy that requires physical engagement the interface is aimed at emphasizing movement, timing, emotion and fun rather than functionality and efficiency. Thus, it promotes physical activity and pleasure over complexity of interaction. Unlike dance motions, which are comparable to some degree to using a Hula Hoop, Hula Hoop motions are significantly simpler and it does usually not require much training in order to learn how to use it. Furthermore, there is no choreography, music or sound needed to aid or inspire the dancer. Also it was much simpler to collect and interpret data based on Hula Hoop movements compared to more intricate ones. Energy, in this sense, is determined by the participant’s actions, transferred to the system via suitable sensors, which in turn changes the entity’s behaviour in a predefined and consistent way. Last but not least, as forming part of an exhibition, the experience should be entertaining and the concept as self-explanatory as possible.

The scenario thus involved a fairly abstract virtual environment inhabited by a collection of rings that would respond in various ways to the participant. An example of a person inside the interaction space is shown in Figure 3-1 on page 72, and Figure 3-2 illustrates the content of the abstract VE. However, as detailed in Section 3.4.3, the underlying algorithm regulating the behaviour of the VE is carefully designed and governed by two separate rule systems. One describes the general system behaviour while the other lays down the foundations of how user action is translated and affects the VE.

Regarding presence and interaction our intention was to show that interaction tightly affects the experience in terms of realism and that more action and engagement results in far more compelling experiences than others. Also we were interested in generating a consistent and reliable mapping for such an unusual UI that allowed its users to attain some sense of achievement and fulfilment. Note that the WiiFit[6] released in 2008, whose balance board evaluates user mass and balance, is a very similar experience allowing the player to virtually hula hoop (see Figure 3-3 on page 73 for detail). The main difference between the two experiences is that in ours a real Hula Hoop is the interaction device while in the WiiFit game real body movements are used to animate a virtual character that is seemingly hula hooping.

[6] http://www.nintendo.com/wiifit

Figure 3-1. Close-up of a person interacting with the final environment.

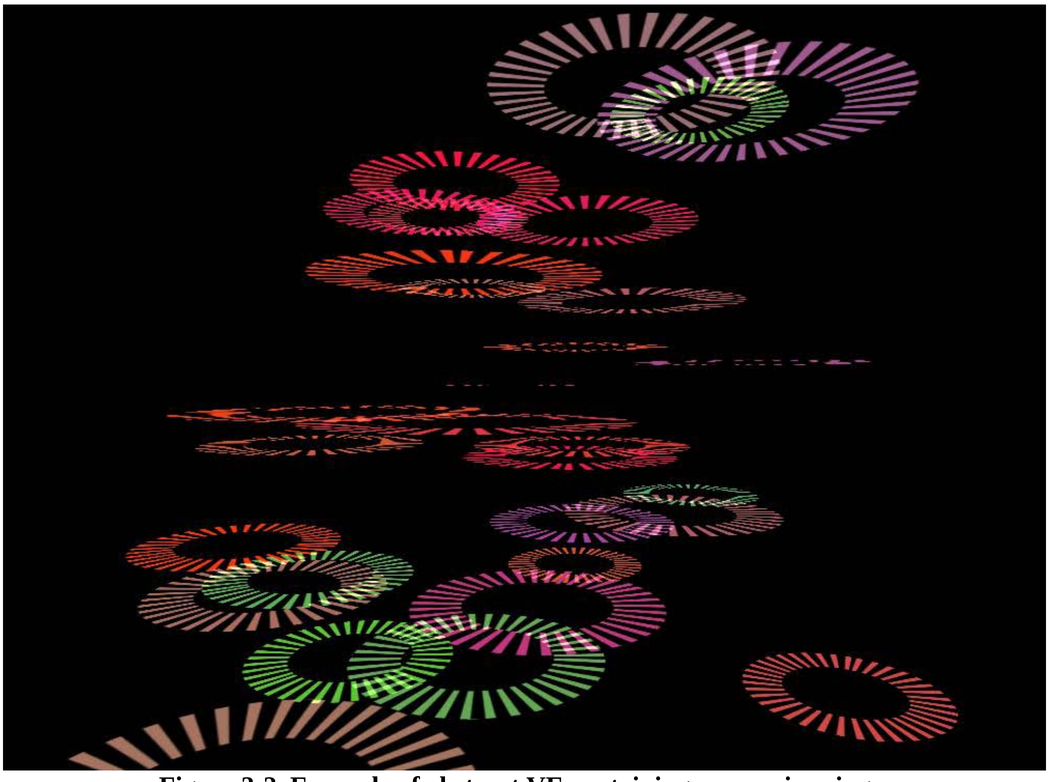

Figure 3-2. Example of abstract VE containing responsive rings.

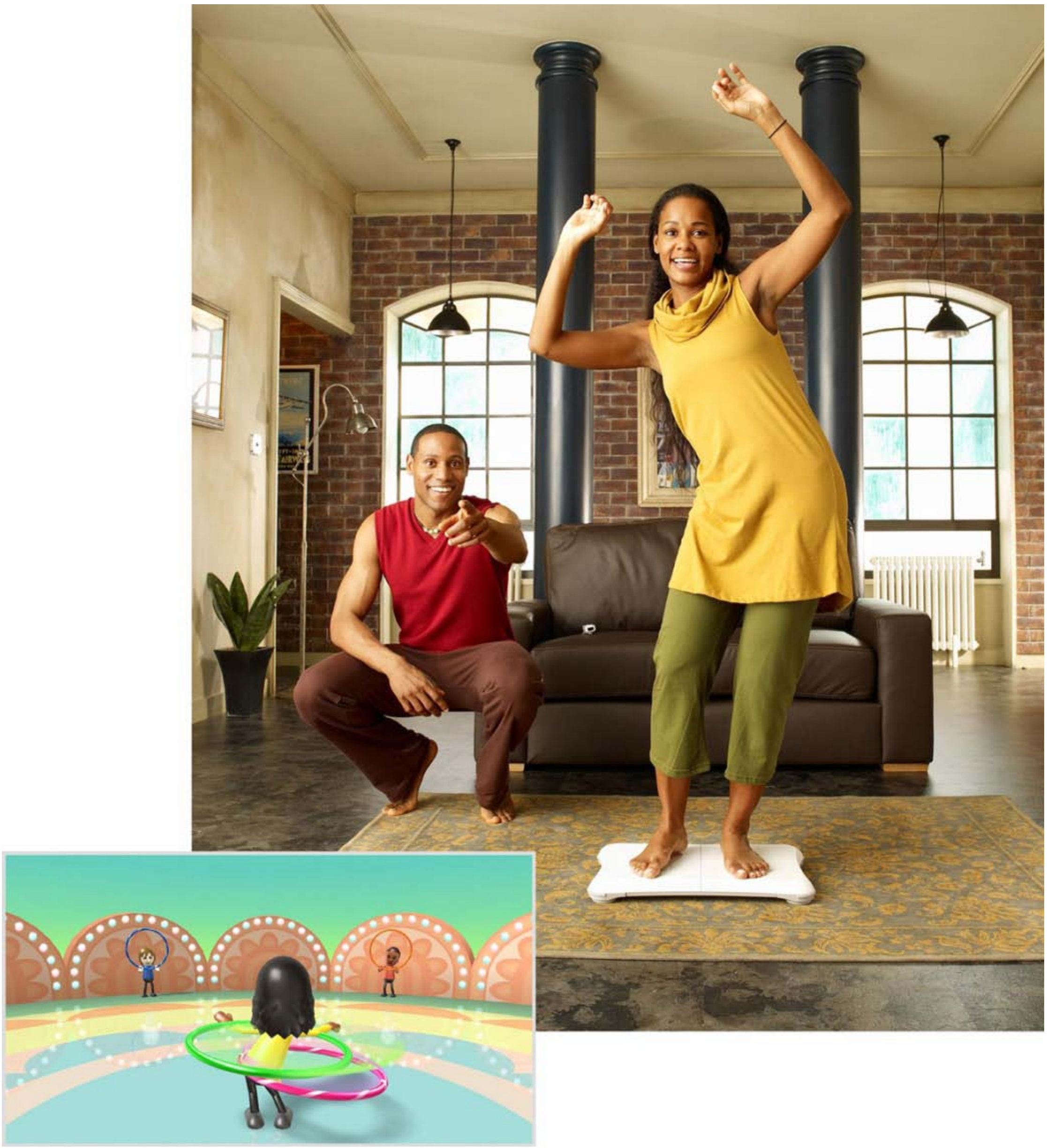

Figure 3-3. A person using the WiiFit to play a game involving a Hula Hoop.

3.3 Experimental Aims and Expectations

The basic hypothesis of this work was that the greater engagement the visitor would display (to the environment) the more active and motivating itself the environment would become in return, forming a positive feedback loop originating in the participant's actions. Furthermore, this would result in the visitor interpreting parts or all of the VE as real as it would respond to them in a consistent manner. Thus a sense of (correlational) presence, or plausibility, could be achieved merely via action. This implies that a greater number or greater use of available Valid Effectual Actions results in enhanced Psi, and one way of testing this is by evaluating two conditions, one in which some VEAs are available and another one where there are none. Note that our focus lies in Psi but not PI, so whatever the impact the lack of VEAs may have on the body and thus possibly on the perception of PI is ignored in this study. Now, if VEAs and sensory variations in the environment can be interpreted in terms of one's own actions, then the result should be Psi.

3.4 Design

3.4.1 Variables

A between-groups design was chosen with the two-level factor being the degree to which subject's subjectively rated their own achievement related to the ease of use of the interaction device. Since there was no objective measure for performance, this variable rather relates to the joyfulness of the experience. The two conditions are listed in

Table 3-1 below. The variables and conditions intend to test the use of SCs in a direct way. While in Condition 2 there are essentially no Valid Sensorimotor Actions because any action has an arbitrary result, Condition 1 on the other hand reflects on a set of VEAs.

Table 3-1. Two-level factor design of the study. Condition 1 has some VEAs enabled while Condition 2 has none available.

Condition 1 (responsive)	**The environment would respond interactively to participant movements.**
Condition 2 (random)	**The environment would appear to respond interactively but really a pre-recorded interaction scheme from a previous visitor was simply looped and used to control the output of the environment.**

3.4.2 Apparatus

Although the study had to be carried out outside the limits of a laboratory we still managed to maintain a substantial amount of control through our setup.

The VE displayed audiovisual information and since there was no powerwall or other VR equipment available we decided to build our own passive stereo display using a projection screen measuring 3-by-2.5 metres, a set of twin projectors, and a collection of passive stereo glasses. Due to space constraints we opted for a suspended front projection. Section 3.4.3 outlines the main development steps of a standard technique for developing a passive stereo display using equipment such as the above. The VE software was written in C++ and OpenGL[7].

Regarding the audio display we used a set of speakers including subwoofer. The subwoofer was mounted behind the projection screen while the speakers where suspended at its top-left and right corners. We tested some other configurations, and although there was no spatial sound involved, we found this to be the best arrangement.

As a sensor to monitor the space and visitors' actions we used a standard 1.3 megapixel webcam with a custom-built infrared filter so that most but not all of the visible spectrum would be eliminated and only the infrared band was captured by the CCD. The camera was suspended above the interaction space, which was a dedicated space about 2m in diameter and roughly 1.5 metres in front of the projection screen. The software for monitoring the space, tracking and processing people's and their Hula Hoops' movements was aided by OpenCV[8], a high-level open-source library designed for continuous image capturing and image processing [Bradski and Kaehler, 2008]. We also used a special infrared light source suspended also from the ceiling and covering most of the interaction space.

---

[7] http://www.opengl.org
[8] http://opencv.willowgarage.com/wiki/, http://sourceforge.net/projects/opencvlibrary/

As the exhibition space had very large windows and because the exhibition itself took place during the daytime our aim was to eliminate as much daylight as possible, also with regards to the passive stereo display. For that purpose we designed and manufactured a suspended canopy frame construction that would block out even more sunlight. The canopy was made of black felt. A sketch of the frame, the passive projection display, projector setup and parts of the exhibition space are shown, to scale, in Figure 3-4 below. Finally, our interaction device was a standard off-the-shelf Hula Hoop which didn't require any preparation or handling prior to the exhibition opening.

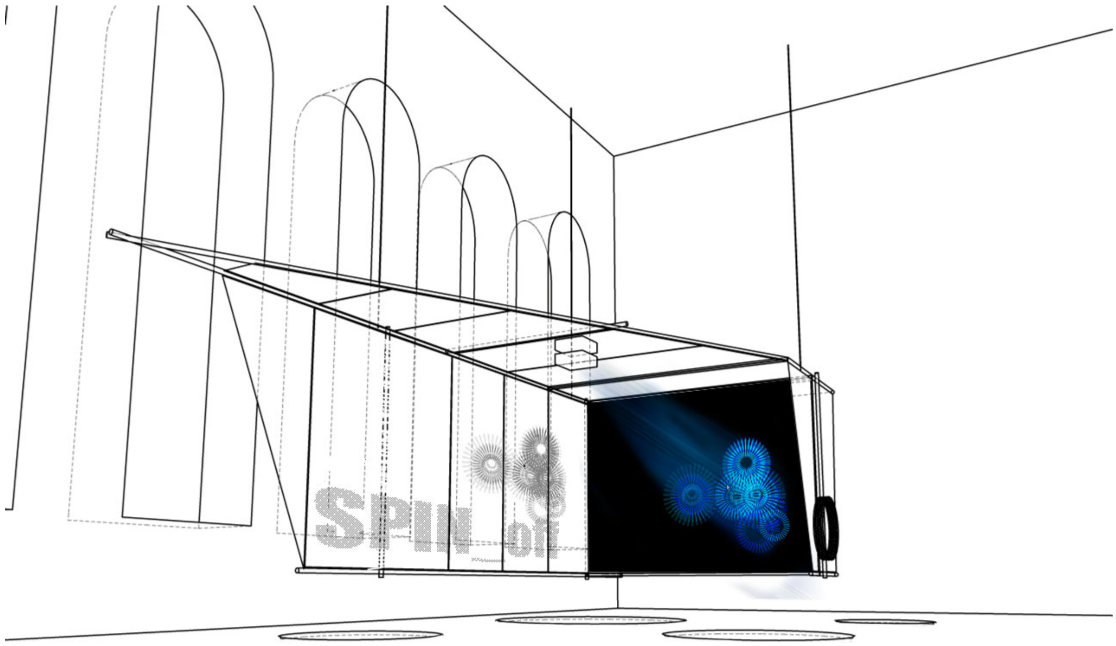

Figure 3-4. Sketch (to scale) of canopy frame, passive projection setup and the windowed side of the exhibition space.

3.4.3 Software Implementation I: Virtual Environment

There were two distinct software components that needed to be developed and set up to exchange data with one another. One part consisted of the virtual environment where the environment and its inhabitants should display a minimum level of intelligence or intelligibility in response to some input vector defined by some features identified in the movements of a participant when using the Hula Hoop.

The agents' behaviours were controlled by a computer model of coordinated animation called boids [Reynolds, 1987]. It has been used in many applications to simulate the behaviour of animals such as bird flocks or fish schools and represents a very simple but visually powerful way of modelling complex behaviour. The boids flocking model describes how a collection of animated objects can display pseudo-intelligent behaviour using only three rules which affect their individual motions. The rules are summarized in Table 3-2:

Table 3-2. Three flocking rules to determine individual boid behaviour.

Separation	**Steer to avoid crowding local flock mates.**
Alignment	**Steer towards the average heading of local flock mates.**
Cohesion	**Steer to move to the average position of local flock mates.**

Each boid object can in theory move freely within the entire environment; however the three flocking rules force it to react to its immediate neighbourhood rather than events occurring in the entire environment or even the entire system of boids. So the behaviour of a single boid is only dependent on those other boids that are within its immediate surroundings. This neighbourhood is defined by Euclidean distance plus the angle relative to the boid's direction of movement. These simple rules were used to induce a basic movement pattern into the agents. Programmatically, it is fairly simple to implement such a system and one can easily add more complex individual behaviour on top of these rules. Attractors, for example, ensure that the boids are attracted to them forcing them to move about within some limited space.

Since our interaction metaphor was that of energy transfer between the real and virtual world we devised a simple and comprehensive set of four energy states that would regulate the energy level and either raise or lower it. The differences between them should be graphically reflected by changing behaviour of the boids system such that it is easy to discriminate between each stage and this is precisely why we chose to use a simple but flexible and modular system such as the boids one. The four energy levels are outlined in Table 3-3:

Table 3-3. Overview of the four energy states used for controlling the agents.

1. Idle	**No one is inside the action space. The boids behaviour is completely determined by the internal rule system and no exterior energy is added.**
2. Person present	**A person has entered the action space although without a Hula Hoop. The boids acknowledge the presence of the user and consequently the energy level is raised and they become more agitated.**
3. Person and/or Hula Hoop present	**The global energy level is raised even more if a Hula Hoop is present regardless of there being a person or not. This is to represent the system's excitement and acts as an incentive to further engage with the system.**
4. Person is Hula Hooping	**The highest level of energy transfer and resulting excitation of the boid system is reached by hula-hooping.**

Now, the basic target for the boids was to swirl around a fixed axis in virtual space. Any person inside the interaction space thus initially had the impression that a swarm of virtual agents was moving in front of them. Aside from the boids rules, the agents had several parameters that could be individually tweaked: colour, spin speed, spin direction, sound and movement radius. Colour refers to the object's colour and spin speed and direction refers to the direction and speed of rotation about its own axis. A second movement factor, movement radius determined the radius within it would move about the focal point. Every boid had a different spin rotation and speed and the movement radius also varied over time. Greater excitation would generally result in greater radii. Figure 3-5 explains the difference between the two. While the overall and global movement of each and every agent was determined by the flocking rules, we included another parameter relating somewhat to their spinning movement. If the spinning movement was zero, the agent would simply spin around its own axis, but any non-zero value would make it rotate around an imaginary axis and the greater the radius the greater the distance to that axis. At the same time however, the agent adhered to the global rule system.

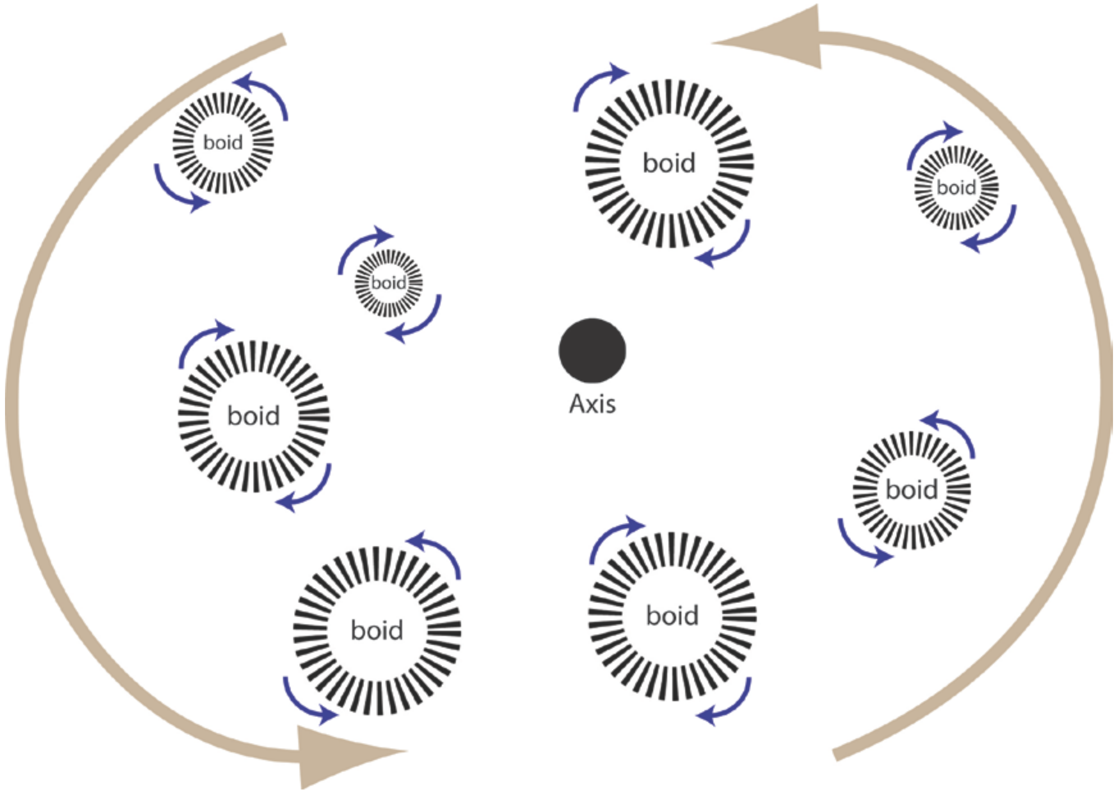

Figure 3-5. Individual and flock movement. Large arrows refer to the global rotations of all boids about the central axis and small arrows to individual spin directions (and speed). Both measures can be altered depending on the energy state and individual boids are assigned a mean value for both around which they are allowed fluctuate.

A lower energy state affects all of those parameters at a more uniform rate so that the majority of the boids would have the same or similar behavioural pattern. Some examples are illustrated in Figure 3-6, Figure 3-7 and some pictures from the entire installation space is shown in Figure 3-8 and Figure 3-9 on page 81.

In reality higher energy levels produce an increased rate of individuality, or randomness among the virtual population. In this sense, each agent had an individual level of excitation while the global level of excitation reflects the

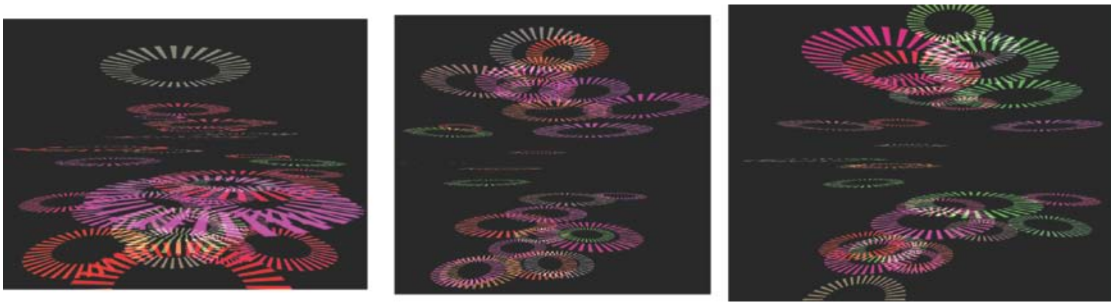

Figure 3-6. Sequence of boid behaviour showing various degrees of higher energy states.

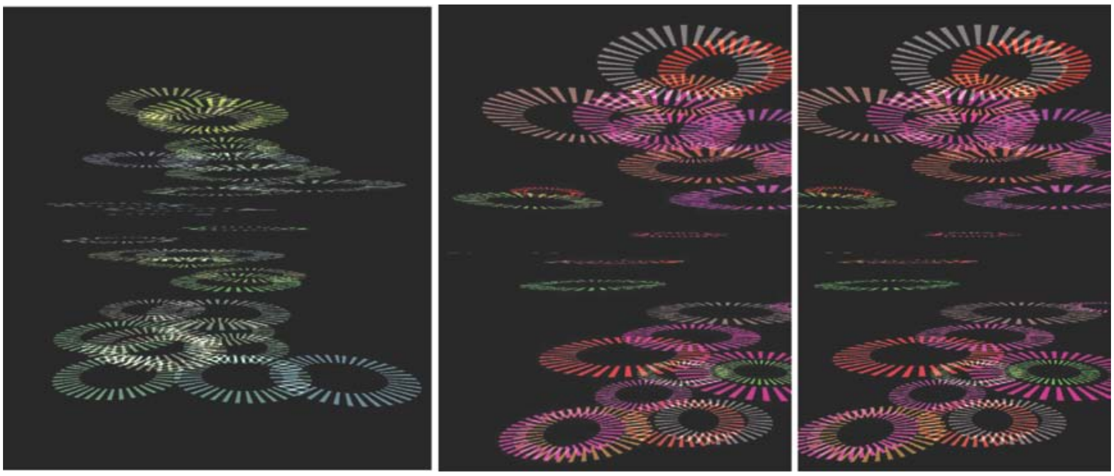

Figure 3-7. Sequence of boid behaviour. Left-hand image depicts a low-energy state while the two right-hand images show a higher energy state.

current energy level and the sum total of the agents' individual energy levels would, except for the highest energy level, never exceed its actual state (1, 2, or 3). The fourth energy level presented a special case in this interpretation because it did allow the sum-total to exceed 4 by any degree. This put emphasis on the fourth energy level where potentially the most reactions were unleashed. In order to sustain and benefit from this energy level one had to hula hoop for a prolonged period. Initially, therefore, all individual energy states were set to the same value according to the simple equation given in Equation 3-1 below and were only dependent on the number of agents. Hence, the global energy level was initially 1.

$$\sum_{i=1}^{n} \varepsilon_i = \frac{1}{n}$$

(Equation 3-1),

where ε refers to the energy level and n is the number of agents in the environment.

Sound was another important aspect of the system and each agent would emit a continuous spatial sound sample relating to its 3D position in space whose frequency would change depending on the agent's individual excitation.

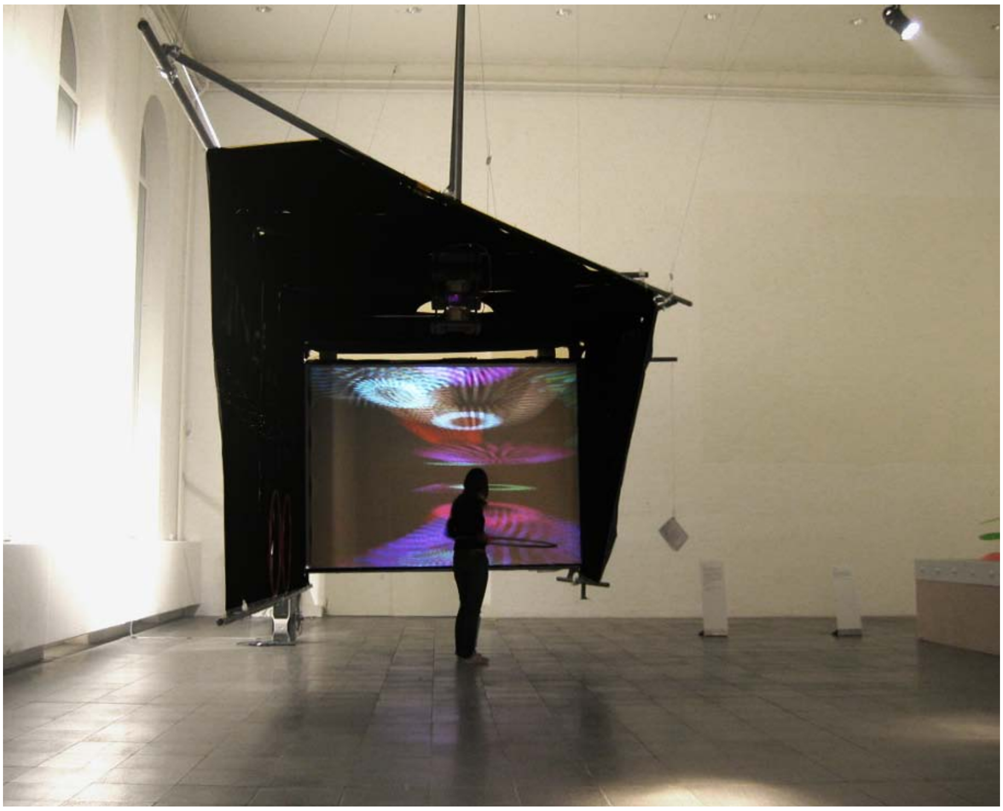

Figure 3-8. Frontal view of the installation with person interacting.

Figure 3-9. Side view of the installation space with person interacting.

3.4.4 Software Implementation II: Tracking the Real Space

3.4.4.1 Design Considerations

As outlined in Section 0 we used some functions contained in the OpenCV library in order to help us with the image processing and tracking of objects. Initially, our goal was to leave open the possibility of having more than one user affect the VE. One of the problems that we encountered during trials was that the camera lens was not powerful enough to track a space large enough to fit more than one person. Thus the actual implementation was single-user only. The space inside the canopy, however, was much larger than the interaction space surveyed by the camera and thus more than one person could be present in that space at the same time. As a consequence people's personal interaction spaces would often overlap though only one person would be tracked at a given time. Regarding the data captured by the suspended camera, this would often result in problems because when there was more than one Hula Hoop visible, the OpenCV routines would often skip back and forth between one and the other resulting in sudden changes of energy states even though all people involved were using appropriate Hula Hoop movements. This was caused mainly by changing lighting conditions due to shadows introduced by the participants. While one person casts fairly predictable types of umbrae onto the floor they may overlap with two or more people resulting in completely different compositions. It is a hard problem to remove shadows from real-time imagery so we did not tackle it and concentrated our efforts on identifying the number and positions of people and Hula Hoops, track them, if possible, and also attempt to capture some lower-degree moments relating to motion of the Hula Hoop and its velocity. The data is then vectorized and passed on to the particle system controlling the behaviour of the individual agents. Even though we did not use any markers for tracking, due to its regular shape the Hula Hoop could itself be considered as a marker.

The following sub-section outlines the major steps involved.

3.4.4.2 The Generalised Hough Transform

One algorithm that excels at finding and fitting arbitrary shapes including circular ones (like a Hula Hoop) in a noisy image is called the (generalised) Hough transform [Duda and Hart, 1972, Ballard, 1981]. The Hough transform was originally concerned with reconstructing imperfect lines in an image but was later extended to include arbitrary shapes, although in practice it is mostly used to find lines and circles. Candidate shapes are detected by using a procedure that is akin to multi-dimensional histogramming. Briefly, the algorithm represents lines in image space as points in Hough space by rearranging its parameterized function (see Equation 3-2 and 3-3).

$$y = \left(-\frac{\cos\theta}{\sin\theta} \right) x + \left(\frac{r}{\sin\theta} \right)$$ **(Equation 3-2)**

$$r = x\cos\theta + y\sin\theta$$ **(Equation 3-3)**

where x and y are Cartesian coordinates and r and θ are polar coordinates. A line in a coordinate system hence becomes a point in polar coordinates because every point on the line described in Equation 3-3 is only dependent on the specific (r,θ) combination. Every point (x, y) in the image plane that lies on line a satisfies the equation with (r_a, θ_a), see Figure 3-10 for details. The greater the number of points that are clustered around a point (r,θ) in Hough space the more likely therefore the probability that a visible but degraded line exists in the original image. Point coordinates (x, y) are known while the correct parameters (r,θ) are initially unknown.

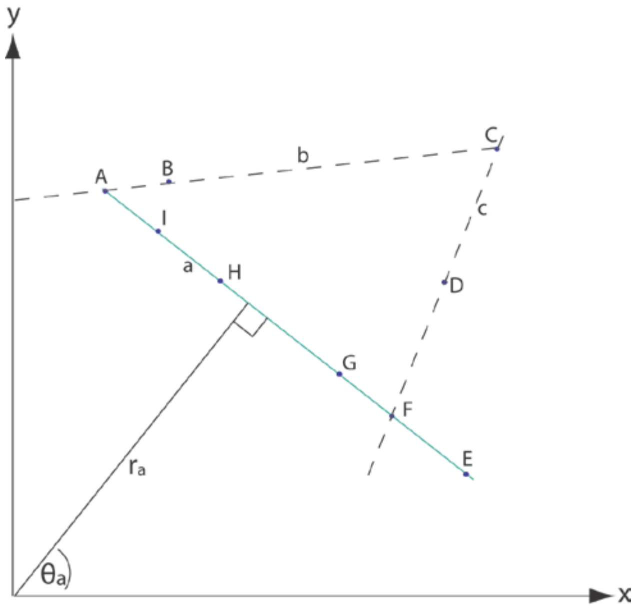

Figure 3-10. Given a number of points A through I, we can generate three lines with at least three of the points on them – in reality the Hough Transform evaluates many more possibilities. In Hough space the axes are not determined by (x, y) but in terms of the (r, θ) parameters the line.

For each line in image space we can now associate a unique point (r, θ) in Hough space and for an imperfect line that has been degraded due to noise or pre-processing, the most likely line connecting the largest number of points will be the point in the Hough image that receives the most "votes". Remember that a point in Hough space really is a line in image space.

The algorithm effectively fits a large number of lines in the original image and checks which points lie on it. In our example above points $AIHGFE$ are collinear and lie on line a, while ABC lie on second line b and CDF on a third line c. Parametrizing the three lines they become points in Hough space and for each of them we check how many "votes" they receive. Clearly line a receives more than the other two (i.e. six over three). If we only reconstruct a single line we therefore choose line a over the other two and points B, C and D are effectively treated as noise.

In order to detect more general image features a similar transform exists for other shapes if they can be parameterized into the general form presented in Equation 3-4.

$$f\left(a_1, a_2, ..., a_3, x, y\right) = 0$$ **(Equation 3-4)**

A circle can be represented by three parameters as in Equation 3-5. The search space therefore becomes three-dimensional but the algorithm essentially works in the same way as the original 2D transform described above.

$$(x-a)^2 + (y-b)^2 - r^2 = 0$$ **(Equation 3-5)**

The input image requires some pre-processing such as edge detection (e.g. Sobel filter). The Hough transform is a very efficient way to detect simple geometric shapes in noisy images and we chose it to detect the occurrence and location of Hula Hoops in a captured frame.

3.4.4.3 Hula Hoop Detection and Tracking Algorithm

The remainder algorithm is straightforward as most processes are included in the OpenCV library and uses standard image processing techniques [Foley et al., 1995, Sonka et al., 2007]. Procedures such as continuous image capture, simple pre-processing techniques including smoothing and edge detection can be done quite efficiently using OpenCV. To facilitate thresholding a running average is accumulated and subtracted from every new frame. A typical camera sequence of the interaction space is shown in Figure 3-11.

All non-zero pixels in the resulting binary image are summed and if the number is higher than a threshold there is at least one large object in the image which does not result from sudden changes in lighting or other external factors. This is a fairly safe assumption to make at this stage since the running average and thresholding ensure that only spontaneous and sudden changes appear as white areas in the thresholded image and therefore small but continuous changes in the lighting conditions due to erroneous and changing sunlight entering the scene are either removed by this technique or do not impair the Hough circle detection algorithm. Results of background subtraction and thresholding are shown in Figure 3-12 and Figure 3-13, respectively.

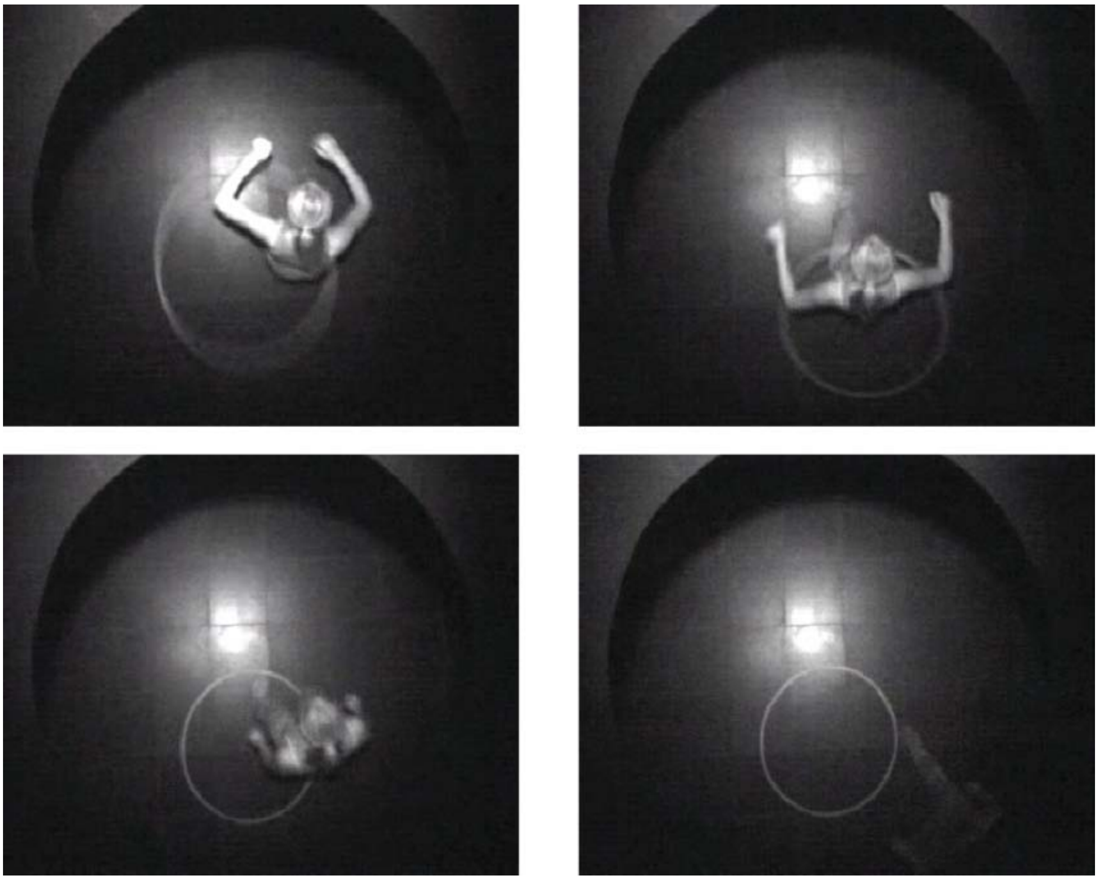

Figure 3-11. Four-image sequence (from top-left to bottom-right) of a woman using a Hula Hoop. Note that the last image in the sequence (i.e. bottom-right) the Hula Hoop has fallen on the floor.

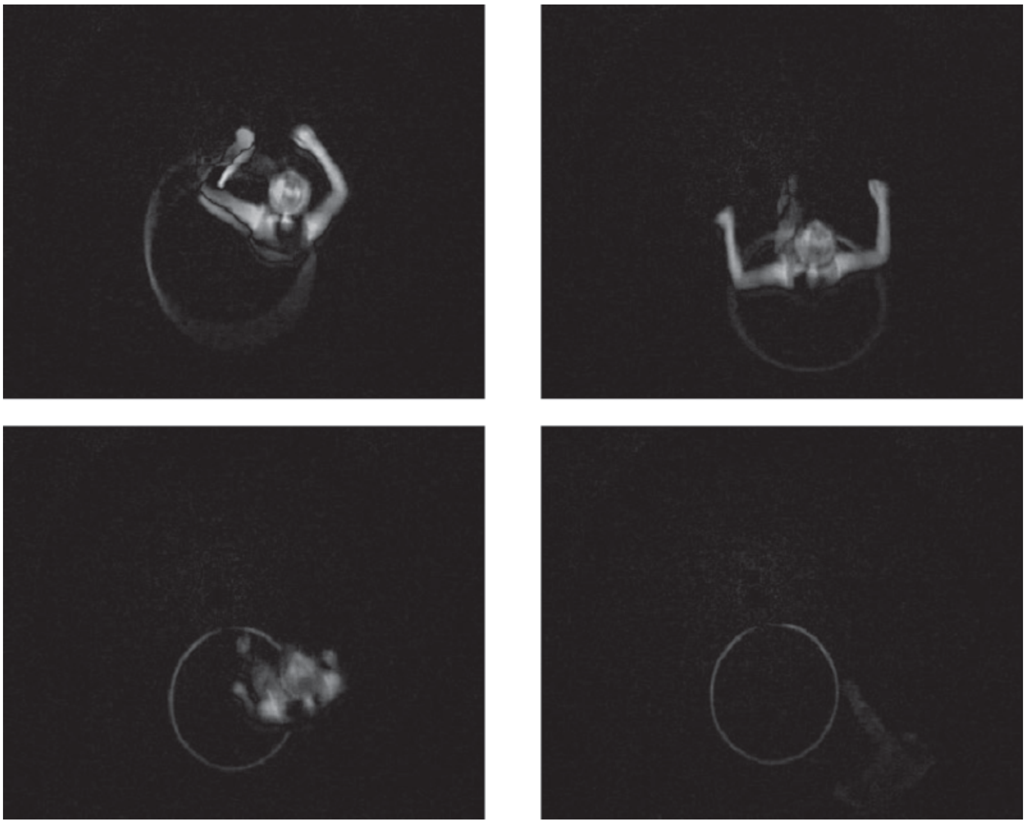

Figure 3-12. Greyscale images resulting from background subtraction with current frame and running average. Note the image is not binary but still contains 8 bits of information per pixel.

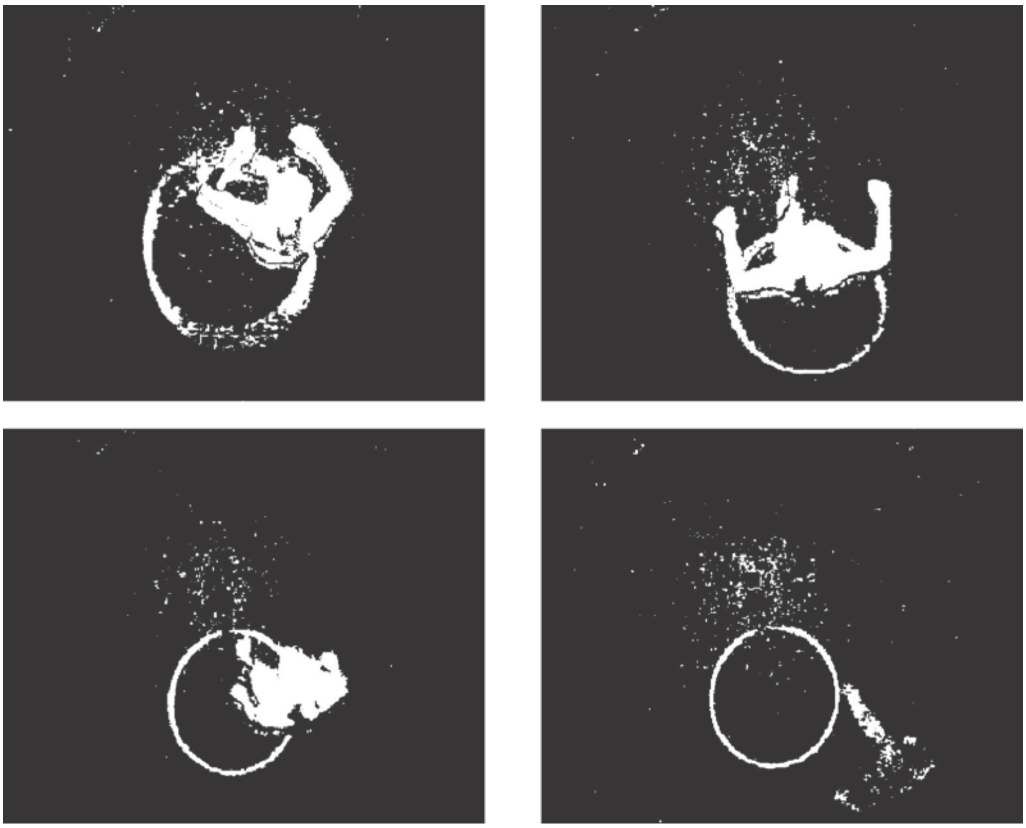

Figure 3-13. Thresholded binary image sequence resulting from the background subtracted imagery in Figure 3-12. Note the occurrence of noise due to light reflections incident on the floor interacting with the person.

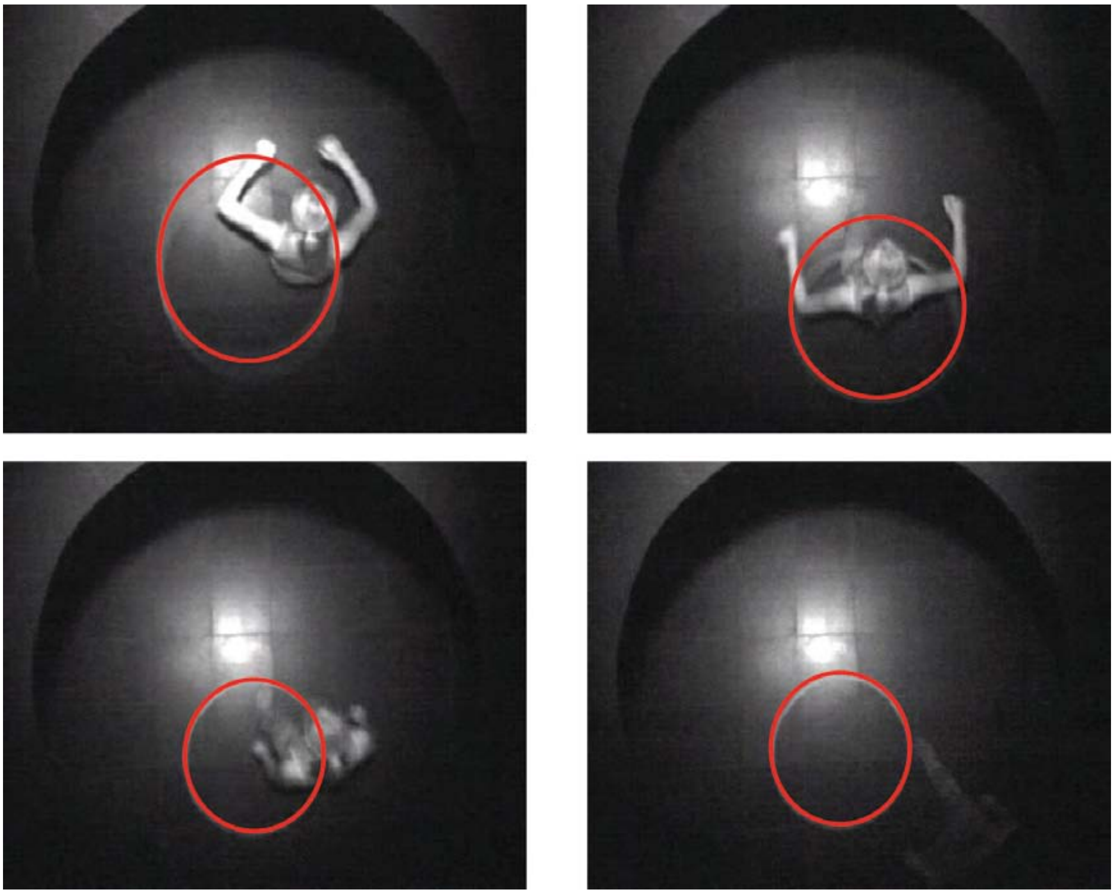

Figure 3-14. Detected Hough circles added into the original frame sequence. The algorithm does not always yield exact results as in the top-right frame for example. Instead it always yields the best fit given the data. Note that the algorithm is very robust although noise is present in the input images (c.f. Figure 3-13).

We construct a search region around the non-zero pixels in the threshold image and apply a simple Mean Shift algorithm which is normally used to find the object centre from its back projection. For binary images, however, it can also be used as a simple method to track an object over a sequence of several continuous frames [Fukunaga and Hostetler, 1975]. We then attempt to find circular shapes using the Hough transform described in Section 3.4.4.2. If it does not detect any circles we assume that only a person is present in the interaction space. If there is a circle, we assume that it is a Hula Hoop and not noise, calculate its velocity and also keep track of the duration since its first occurrence and since its last non-zero velocity. The latter is important since little or no velocity implies that the Hula Hoop is not moving and therefore not operated at all or incorrectly by a participant (i.e. it has fallen on the floor). The detected circles for our example sequence are shown in Figure 3-14, combined with the original unprocessed imagery.

Now, the current position Hula Hoop is tracked using the Mean Shift algorithm. Since it operates using an initial search window and once it has been tracked for several frames, any newly occurring objects such as Hula Hoops, people or people with Hula Hoops are essentially ignored by this two-stage algorithm. Given the task the whole procedure is therefore very robust and noise-insensitive.

3.4.4.4 Interpreting the Data

In Table 3-3 on page 78 we showed the four different energy states and their effects on the environment and in the previous section we discussed how data on inferred body movements were gathered and analyzed. Table 3-4 shows how the actual movements are translated into one of the energy states. The mapping between parameters is very simple and one-dimensional only but the resulting environment is very diverse (cf. Figure 3-6 and Figure 3-7 on page 80).

Table 3-4. Relationship between parameters and detected objects in the interaction space and the four energy states.

1. Idle	**No changes compared to background image are detected by the algorithm.**
2. Person present	**The number of non-zero pixels in the thresholded image is greater than some predefined threshold.**
3. Person and/or Hula Hoop present	**Same as 2 AND the Hough transform has returned at least one circle whose radius is greater than some *r*.**
4. Person is Hula Hooping	**Same as 3 AND the person has been using Hula Hoop for some time (i.e. continuous non-zero velocity).**

3.4.5 Task

Given the nature of the exploratory environment which depended largely on the actions of the individual participants it was difficult to formulate a formal task. Hence the task the participants were given was:

> "*Use the Hula Hoop and your body in order to interact with the environment and its inhabitants and focus on how you can affect them.*"

They were not given an explicit time limit. No further mention was made of the elements of the VE or how participants could make use of the Hula Hoop – even though the latter point is essentially implicit, it was by no means obvious (or intended) that they would use it in the expected way. Earlier observations showed that many people who were unable to Hula Hoop tried to elicit a response from the system by holding it in their hands and moving it around the interaction space.

3.4.6 Population

We conducted a study with 22 (8 male and 14 female) volunteering visitors of the exhibition in two groups (experimental group A and control group B) that were randomly assigned. Since the study was conducted during the exhibition, participants were ordinary visitors who we invited to take part in our study. Participation in the study was purely voluntary and participants did not receive financial compensation for their taking part in the study. The study was carried out in one day only.

3.4.7 Procedure

The procedure was divided into three stages. In an introductory period the experimenters familiarised the participants with the use of the passive stereo glasses. The participants also gave signed consent that they had been given sufficient information about the study. The consent form is printed in Appendix D. The participants were randomly assigned to either group A or group B with the goal of balancing both groups, which eventually contained 11 subjects each.

The second stage consisted of the actual experiment, where each participant entered the interaction space alone, equipped with a Hula Hoop. As the experiments were carried out while the exhibition was open the experimenter ensured that no other visitor would interrupt or distract the participant. The experimenter also timed each experiment and noted down the result.

In the post-experiment phase each participant filled in a questionnaire comprised of twenty questions, some inquiring basic demographic information. The larger part of the form however dealt with subjective performance and we used a slightly modified of the questionnaire by Witmer and Singer [Witmer and Singer, 1998] in order to assess to assess the quality of the experience and the interaction. The questionnaire also contained a section where the participant could comment freely. It is printed in Appendix A.

Experimental group A was presented with the normal interactive environment where correct and constant use of the Hula Hoop would attract the virtual agents and alter their energy state resulting in more energetic movement and behaviour. In the control group B however, any effort made by participant had no effect on the development of the boids system and was therefore not interactive whatsoever. Rather, the system's behaviour was determined by data that was recorded prior to the experiment.

3.4.8 Response Variables

We considered three different variables in the analysis, including *control* over the environment, *impact* over one's own actions and correlational presence (*identification* with the environment). The variables were constructed from different questionnaire items, each on a 7-point Likert-scale with the scores adjusted for the analysis.

3.5 Results

We assigned each answer of the 20-piece questionnaire printed in Appendix A into one of three variables. These were *control*, *identification* with the VE/objects and action *impact*. The mean scores for both conditions are presented in Figure 3-15 below. The scores show that for all three variables responses were higher in the interactive group A compared to random group B. For the interactive condition the mean score was 4.1 and for the random condition it was 3.0.

We carried out an analysis using normal Analysis of Covariance. There was a significant difference at the 10% level but not at 5% between the groups in their perceived sense of *control* over the environment and *impact* of one's own actions, where those in the interactive group A had the higher mean questionnaire scores. There was also a significant difference at the 5% level between the groups with respect to the perceived *identification* with the VE: Results from group A showed that the more the output of the VE was meaningfully attributed to one's own body movements (which was lacking for group B since the system's response was random), the more likely one was to associate with it. In addition the perceived time spent in the environment was also positively and significantly correlated with this response variable. In both cases a Bera-Jarque test did not reject the hypothesis of normality of the residual errors.

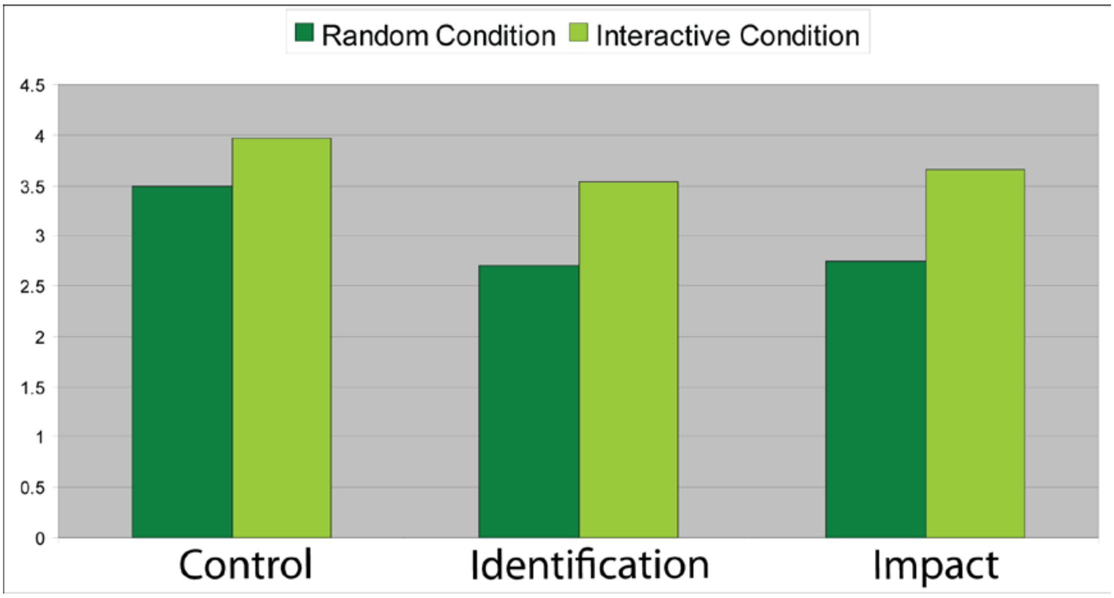

Figure 3-15. Mean scores for the three variables control, identification and impact.

We were thus able to confirm one interesting aspect regarding presence, namely that one necessary condition for plausibility is consistency of the stream of sensory input and the expected outcome [Gillies and Slater, 2005].

3.6 Chapter Summary

In this chapter we presented a MR environment that allows participants to interact with virtual entities using a real existing toy – a Hula Hoop. The virtual environment consists of a set of boids that display emergent behaviour and react toward the actions of the participant in four distinct stages.

We employed a hula hoop as an interaction device for a virtual environment thereby merging two different domains: a real space that is characterized by a physical interface that is not custom-made but exploited for use with VEs.

We also showed how a very simple rule-based algorithm can be tuned to yield a group of responses that are discernible by the human participant. Until now and to the best of our knowledge, boids systems have only been used statically in the sense that the internal parameters of the algorithm would not change throughout a simulation. While attractors and other objects can occur at random points and affect the behaviour of the boids system these items do not change their internal behavioural rule system. Our system successfully demonstrates that the rules of the system can be changed by an interactive component – in this application the energy that is provided by the participant.

We demonstrated that using a hula hoop in a real space can yield a set of “commands” for a computer interface, which is not only interesting for interaction designers regarding interface design and interaction metaphors, it also stresses the simplicity of the interface itself. We have created a playful interaction space that aims to reduce the gap between real and virtual spaces in an entertaining fashion. We were also able to verify a vital requirement of presence, namely consistency between the streams of sensory input.

Many hundred visitors enjoyed the installation during the exhibition in Copenhagen in November 2006.

4 Implicit Physiological Interaction and Plausibility – The Physiological Mirror

4.1 Introduction

In this Chapter we introduce a more indirect way of communicating with a VE. While the previous and following chapter deal with very direct means of translating a human action (i.e. thought and body movements) into a meaningful virtual one, the experiment described in this chapter shows how even unconscious and, to some extent, uncontrolled processes can be used for interaction. Moreover, we attempt to enhance the sense of realism of a VE by feeding and visualizing unconscious (physiological) information into the system thus intending to create a stronger relationship between the VE and its human visitor.

As we described in Section 2.4, in a VE there are two dimensions of 'presence' that can be addressed. The first refers to 'place illusion' (PI) the illusion of being in the place depicted by the VE scenario. This can be achieved through the participant being able to perceive the world through normal sensorimotor contingencies. However, even if the participant has the illusion of 'being there', there is a separate question which is concerned with how real the depicted events are perceived: Is there the illusion that events occurring in the VE are really happening? This is referred to as 'plausibility' (Psi) and we introduced the term in Chapter 2. Our general hypothesis is that amongst other things, Psi is enhanced when there are multisensory correlations between the body state and dynamics of the participant and events in the environment. For example, as the person moves forward towards a virtual character, the character somehow responds [Gillies and Slater, 2005, Pan et al., 2008]. If participants see a virtual reflection of themselves in a mirror, that apparently moves as they do, then they will experience this reflection as somehow a reflection of themselves. There are many possibilities of such multisensory correlations. An excellent example is given by the rubber hand illusion, which has been demonstrated also in VR [Slater et al., 2008]. Here, correlation between the visual appearance of touch on a virtual arm and actual touch on the participant's unseen real arm, gives the person the illusion that the virtual arm being apparently touched is owned by them.

In this research we investigate whether the idea of multisensory correlations can also apply to the (unconscious) state of the participant's body. Suppose that changes in the environment reflect changes in the physiological state of the participant – their breathing, their heart activity, and their level of arousal. Would this enhance the probability of the illusion that what they are experiencing is 'really happening'? Essentially, this project can be regarded as a physiological mirror where real physiological input is displayed in a VE as natural as possible. While accurately visualizing some physiological processes is difficult (e.g. heart rate) we chose to express them metaphorically.

We tested this in an experiment using an immersive setup. The scenario consisted of a virtual waiting room of a doctor's office, which was inhabited by a number of seated virtual characters that displayed autonomous behaviours and movements. However, the movements of one of the avatars was determined by the underlying physiological state of the participant, although, as explained below, the movements generated for the autonomous and user-controlled characters were similar and it required a certain degree of focus and self-awareness in order to distinguish between these two streams of events.

Our main goal was to understand if and under what conditions a person can detect this particular real part relating back to their own physiological state in the VE. In this sense, this project bears a lot of similarities to Spin_off presented in Chapter 3, in which we established that correlations between physical movements of one's own body and reactions in an (abstract) environment enhance plausibility and the degree to which a participant determines the meaning and the possible implications between his actions and the reaction of the VE. Unlike Spin_off however, this experiment mainly focuses on mainly activities of which a person is normally not consciously aware.

4.2 *Aims and Expectations*

Our goal was to evaluate the benefits of unconscious interactions towards enhancing one dimension of presence, namely plausibility, and actions were derived from physiological measurements. In particular, we measured heart rate, respiration and skin conductance.

The underlying hypotheses in this project are: Can the unconscious physiological human states be captured, interpreted and displayed, via avatar behaviour, in a recognisable way? Secondly, can a human participant successfully identify this real part from a suitable number of impostors whose behaviour was generated by previously recorded physiological data? Finally, if the participant successfully locates the real part and identifies the correct avatar, will he be able to relate this information back to his own behaviour or will he simply perceive it as a behaviour that is superior and more real than the others?

This should have direct consequences for the level and possibly even control of Psi: the “real part” in the VE comes from the observer himself and therefore he should be able to identify with the VE much more than if this correlation was absent. Events in the VE therefore correlate with one’s own actions, which are unconscious in this case.

4.3 *Design*

4.3.1 Outline

We carried out an experiment that tested these hypotheses. The participant was told in advance that they would enter a virtual Waiting Room and that there would be n other characters there with them. The person is not allowed any interaction such as talking with the virtual characters. The characters are programmed to occasionally look at the person, and change facial expression, and other behaviours (see Section 4.3.4 for detail).

Essentially the participant is asked to select the character that they intuitively connect with (cf. Section 4.3.5). All characters exhibit the same types of behaviour. However, one of them (a different one for each participant) is programmed so that its behaviour is correlated with the physiological state of the person. The behaviour is determined by a set of looped animations and other modifications. The mappings are listed in Table 4-1 below.

Plausibility translates into choosing the character that exhibits correlated responses with the participant's physiological response. Clearly we are asking the participant to identify what is 'real' in this environment, so it is a direct behavioural test of our hypothesis. We call the avatar that reflects back the physiological state of the person the 'R-Avatar' (reflective avatar).

Table 4-1. Description of mapping and actual correlations between real physiological input and avatar behaviour.

Description	Physiological Measurement	Avatar Response
Skin colour as a function of relative change in skin conductance.	Skin conductance (SC) and galvanic skin response (GSR)	Skin colour changes in red channel.
Breathing correlates with real breathing of the participant.	Respiration	Direct mapping between real and virtual respiration.
Changes in posture as a function of heart rate.	Heart rate (HR) and heart rate variability (HRV).	Avatar foot tapping correlating rhythmically with current heart rate.

4.3.2 Variables

The single aim of our study was to examine whether a statistically significant number of people were able to select the R-Avatar given the presence of other characters. There were n avatars in a virtual room and $(n-1)$ of them behaved according to $(n-1)$ different sets of pre-recorded physiological measurement chosen to be respiration, skin conductance and heart rate. The R-Avatar responded to the participant's actual physiology.

Suppose that there are N virtual people in the room. Then the probability of choosing the R-Avatar by chance alone is p = 1/N. If there are n independent trials of the experiment, then the probability of choosing at least x correctly by chance is described in equation 4-1 below.

$$P(x,n,p)=\sum_{r=x}^{n}\binom{n}{x}p^{x}(1-p)^{n-x}$$ **(Equation 4-1)**

We want to choose n and p so that this probability is less than or equal to 0.05 (the conventional significance level).

We thus observe a single binary variable and record and accumulate the participants' selections to the experience. Qualitatively, we are also interested in the reasons that motivated participants to select one avatar over another, so we also pose this question after they have made their selection.

4.3.3 Piloting

There were several motivations for piloting this experiment. One reason stems from the fact that we want our non-responsive avatars to display comparable behaviour to the responsive one so a simple way to do this is to use pre-recorded physiological data. During piloting we therefore also recorded – with individual permission – the signals and saved them for later use that would control any of the other avatars.

The second objective was, of course, to test the technical setup and to confirm that all the tools involved communicated correctly, data generated at the required rate and performed as expected. Since we used several instruments and computers to capture and display data, it was paramount to guarantee smooth operation during the experiments.

Half of the pilot experiments were carried out using a head-mounted display (HMD) but due to the low image quality and resolution we finally switched to a powerwall (see Section 4.3.4.1 below for details). Our observations during trials as well as comments from many pilot participants suggested that image quality does play a crucial role in an experiment that is primarily decided by mental correlation of the observed scene. The fundamental process that guarantees this is an accurate and detailed visual representation allowing precise observations to be made. Poor image quality on the other hand does not elicit the desired detail and quality of perception and participants often failed to detect the more subtle changes such as the avatars' facial colour or breathing rhythm. Ultimately, the quality of the HMD was not good enough to provide us with these properties.

We carried out a three pilots with eight, seven and eight volunteers, respectively. In the first pilot, we used eight different avatars to represent the set of eight characters present in the waiting room. Results from this pilot suggested that volunteers would often select an avatar based on their appearance, tending towards selecting an elderly person, and they possibly also favoured avatars who sat in certain positions. One of the eight volunteers selected the right avatar.

In the second and third pilot we thus controlled for appearance and custom animations and only a single avatar was used and visualized and animated six times in six different positions corresponding to five $(n-1)$ different sets of pre-recorded physiological measurements plus the participant's actual physiology. In the second pilot we controlled for varying degrees of realism of avatar appearance and animations. In this pilot, again, one person selected the R-Avatar.

In addition to this, the third pilot comprised playing back a five-minute voice recording of intended to relax the volunteers. There were two motivations for this. One was to slow down respiration and heart rate and the other to make volunteers more conscious of them. In the last pilot none of the volunteers selected the right avatar. We also exaggerated the respiration animation so that it became more obvious (see Section 5.4.3.3.4 for more details).

During the three pilots we also modified the task given to the volunteers. While in the first pilot we told them that only one of the avatars was controlled by a real person over the internet and that their task was to choose the right one, in the second pilot we asked volunteers to select the character that behaved most realistically and that this decision should be made intuitively. For the third pilot we used a similar task stressing the fact that the decision should be made intuitively.

Some of the responses we received during all of the pilots, in particular the first one, led to us to simplify the environment, most notably by using n copies of the same avatar thus cancelling out the effects of appearance. Other volunteers made their decision based on the position, the lighting or windows present in the environment so that we eventually simplified the virtual waiting room scenario to another more neutral setting (cf. Section 4.3.4.4.).

Since there was no observed gain from using copies of one avatar, in our study we returned to the original idea of using different characters instead of copies of a single avatar. Pilot results also led us to reduce the number of avatars to six.

4.3.4 Apparatus

4.3.4.1 Equipment

The VE is displayed on a 3x2m powerwall via two calibrated projectors with a native resolution of 1024 by 768 and 2500 ANSI Lumens. The head is tracked via a six degree of freedom (6DoF) Intersense IS900 motion tracker, attached to a pair of passive stereo glasses that are worn by the participant in order to perceive the scene in full 3D. There was no audio output. The VE was displayed using a PC running Microsoft Windows XP Professional with an Intel Pentium 4 CPU 3.20GHz, 2GB of RAM and an NVIDIA GeForce GTX 260 graphics card[9]. The VR was displayed by using eXtreme Virtual Reality (XVR)[10] [Carrozzino et al., 2005] and the participant could navigate through it by using a 6DOF standard wand with a direction stick and four buttons. During

[9] http://www.nvidia.com
[10] http://www.vrmedia.it

pilots only we additionally made use of a Fakespace Wide5 HMD[11] which has a field of view of ~150° horizontal/~88° vertical and a screen resolution of 1600 x 1200 at 60 Hz. In Section 4.3.4.5 we explain why it was not used for the actual experiments.

For capturing physiological information, we used a g.Mobilab+[12], a wireless multi-purpose biosignal acquisition device that can be attached to a person's belt. The captured data was transferred via Bluetooth to a PC at 256Hz, which is sufficient in order to extract HR, GSR and respiration in real-time. Physiology was captured using additional sensors attached to the device. A g.FLOWsensor was used for capturing respiration, a g.GSRsensor for GSR and a set of five electrodes for bipolar ECG. The captured data was sent to a PC running Microsoft Windows XP Professional with an Intel Xeon CPU E5320 @ 1.86MHz and 3 GB of RAM. We used Matlab and Simulink[13] in order to process and analyze the data and a g.HeartRate Simulink module to filter respiration and ECG signals and to estimate basic HR and HRV. This and the remaining elements are outlined in more detail in Section 4.3.4.3 below. Although we initially used a Nexus-4[14] device for capturing physiology, which due to its smaller size and simpler setup was our preferred choice, we later switched to the g.Mobilab+ for several reasons. First, there was no built-in protocol for automatic and continuous data transfer from the device to a nearby PC and even though we tested several protocols we found none of them compatible with the device and with Simulink. A result of this was that data transfer and rate became unpredictable and thus essentially useless. A more detailed account of the problems encountered and motives for changing the initial approach is given in Section 4.3.4.5. Secondly, although the Nexus-4 device is much simpler, its measurements are also much less robust against noise introduced due to either movements made by the wearer or plain signal noise. This is especially true for the respiration signal. More detail on the implementation of the physiological data processing is given in Section 4.3.4.3.

[11] http://www.fakespacelabs.com/
[12] http://www.gtec.at
[13] http://mathworks.com
[14] http://www.mindmedia.nl/english/nexus4.php

The model of the waiting room and the virtual characters, which were acquired from Axyz-Design[15], were prepared for XVR in 3D Studio Max[16]. More details about preparing the models, illuminating the scene and character animation can be found in Sections 4.3.4.2 and 4.3.4.4.

4.3.4.2 Character Animation

Our design required all virtual characters and most importantly the R-Avatar to be animated interactively. We first had to create a set of animations and texture masks for manipulating texture properties as well as export them for use with the XVR suite. Finally, we had to control these animations from within our software as a function of the incoming physiological data. Heart rate was set to manipulate the speed and frequency of a foot tapping animation while respiration was designed to map directly to the avatar's animated respiration. Since GSR values represent a much more inflectional description of an internal psychological state than the other two, we did not use the values directly to control an animation but rather applied it change the avatar's skin colour. We chose this metaphor because both GSR and blushes somehow relate to a person's state of arousal.

In this section we will outline what techniques were used to create, export and use animations. Most of the procedures presented, however, are standard techniques for character animation [Steed, 2003]. In order to animate a character a temporal sequence of translations and rotations is applied to the underlying bones of a skinned mesh. Bones are modelled to approximate the skeleton of a human body. They form an interconnected hierarchy, in which a transformation applied to a higher-order bone is also performed on all the bones in the sub-hierarchy of that particular bone. Table 4-2 outlines the basic hierarchy of a limb from upper arm to fingers. Transforming for example the Forearm equally affects all bones that are below it, namely {Hand, Finger0, Finger1, Finger2, Finger3, Finger4}. An example of a partial bone hierarchy and the transformation described above is shown in Figure 4-1.

[15] http://www.axyz-design.com
[16] http://www.autodesk.com

Table 4-2. Limb Hierarchy from clavicle to terminal limbs.

```
Upper Arm
  Forearm
    Hand
      Finger0
      Finger1
      Finger2
      Finger3
      Finger4
```

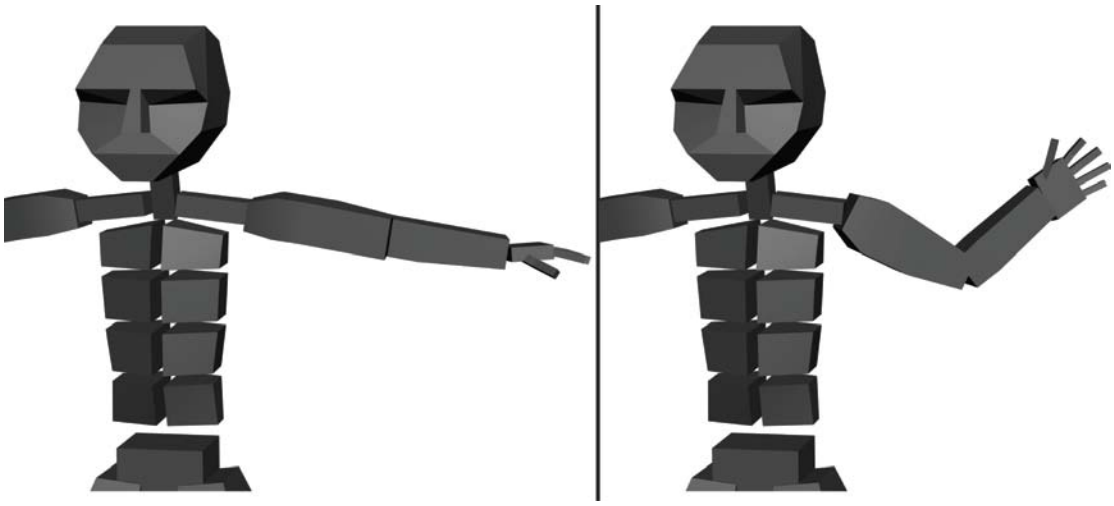

Figure 4-1. Partial bone hierarchy before (left) and after (right) rotation applied to the Forearm. The rotation also affects hand and fingers.

Our design required every character to be seated. In addition, we needed to implement animation for foot tapping of the left and right foot independently as well as continuous breathing. Most of these are fairly straightforward to generate by using the character animation tools in 3D Studio Max. However, obtaining the desired pose and animations for a set of characters is an extremely time-consuming task and a substantial amount of time was spent on preparing ten virtual characters and animations for each of them for use in the XVR environment. Figure 4-2 below shows two perspectives of a virtual character in initial pose and Figure 4-3 shows the same character in a seated posture. Now, while the basic changes in bone configuration from the initial T-pose to a seated posture can be done using motion capture data, often it does not include the correct posture for hand and finger positions and sometimes it occurs that other limbs including head require manual adjustment. Thus, employing a pre-loaded (still) animation such as the seated posture still required us to do extensive post-processing and cleaning for every individual character.

Figure 4-2. Basic pose of a virtual character.

Figure 4-3. Seated and post-processed character.

Once this is done we need to manually prepare breathing animations for each character. We aimed to match the breathing animation of the R-Avatar exactly with the participant's breathing rhythm. Instead of realistic deformations of the mesh around the chest region, our approach involves small rotations applied to the bones in the spine and upper body. Again, this is done individually because, even though the bone hierarchy among the characters should be the same, the meshes can vary to a great extent. Since each mesh is affected by its underlying bone rotations, this can have undesired effects if care is not taken and cause, among other problems, unrealistic overlaps. Therefore, we need to take great care applying rotations to the spine and neck bones. A completed animation is then exported using a Cal3D[17] exporter modified for use with HALCA, a character animation library with an interface to XVR that supports morph animations [Spanlang, 2009]. Morph animations allow us to access an arbitrary slice of the animation by addressing it using a value between 0.0 and 1.0 – its starting and end constellation, respectively – and this system provides an efficient method to control pre-built animations programmatically using HALCA's `setMorph` command.

Finally, we wish to interactively access and modify some portion of the avatar's colour as defined by the texture it is associated with, and our aim is to affect only the facial skin regions. The fundamental idea behind changing an avatar's facial skin colour is to portray a (real) person's level of arousal determined by GSR measurements as a sequence of blushes. The stronger the blush (i.e. greater proportion of red) the greater the person's real arousal. Skin colour, or indeed any arbitrary portion and colour in the character's texture, can be changed programmatically by applying a multiplier to one or more colour channels in a GLSL fragment shader. The Uniform variable that defines the skin colour can be modified at runtime from XVR. We used the alpha channel of the RGBA texture to mask out the region of the texture that we want to modify. Such a mask can easily be prepared using most standard image editing software. See Figure 4-4 for an example where only the face is affected by the colour adjustments in the GLSL fragment shader.

[17] http://home.gna.org/cal3d/

Figure 4-4. Texture for head and eyes of a virtual character (left) and the binary mask outlining exactly the region whose colour should be interactive (i.e. facial skin regions). The black region will be affected by colour changes.

We have now covered all the necessary aspects of character animation (and texturing) that will allow us to control breathing, foot tapping and skin colour interactively – in our experiments based on recorded and real time HR, GSR or respiration data. How we process these is the topic of the following section on Physiological Measurements.

4.3.4.3 Physiological Measurements

4.3.4.3.1 Overview

Our goal was to capture three physiological measurements as well as process and analyze them in real time. These measures are HR, GSR and respiration. As stated in Section 4.3.4.1 we used a g.Mobilab+ to capture the data. The device is shown in Figure 4-5 together with each of the three sensors and an image of a person fully connected to it is shown in Figure 4-6.

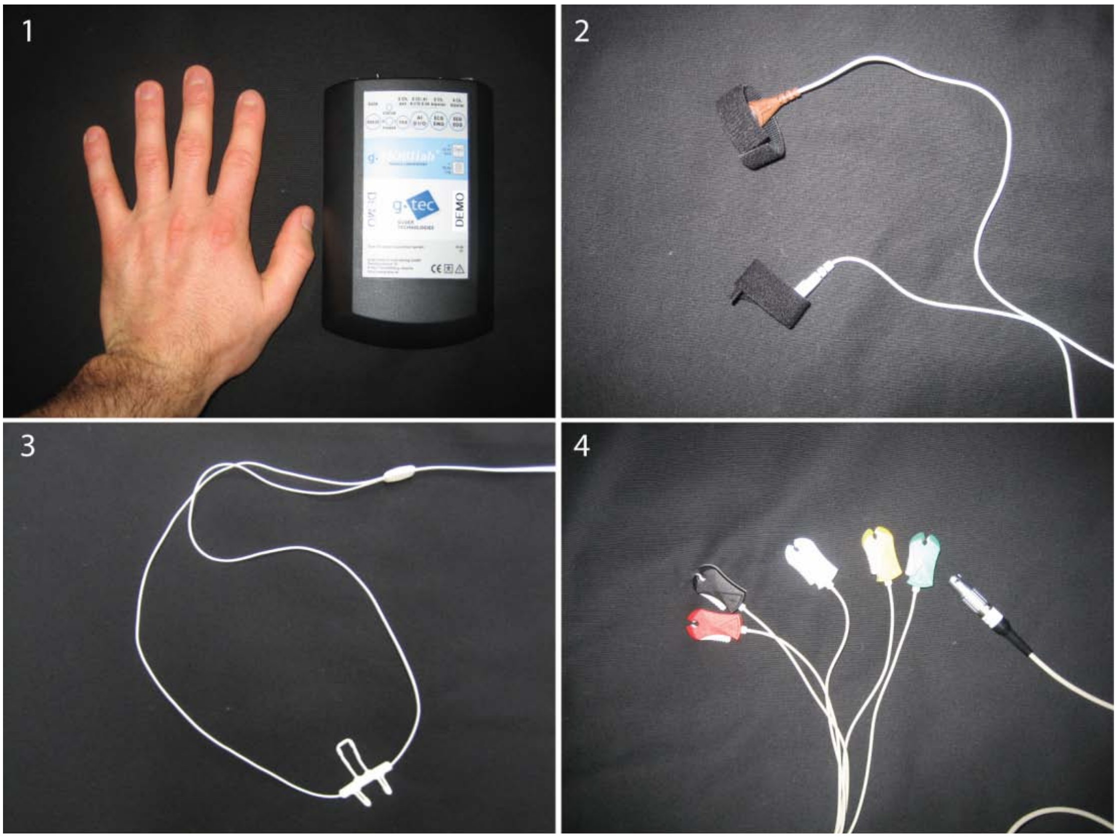

Figure 4-5. (1) Image of the wireless device used for physiological measurements. (2) GSR sensor attached to index and middle fingers, (3) Respiration sensor worn around nose and mouth, (4) ECG sensor, each of the five sensors are attached to various locations across chest and lower arm.

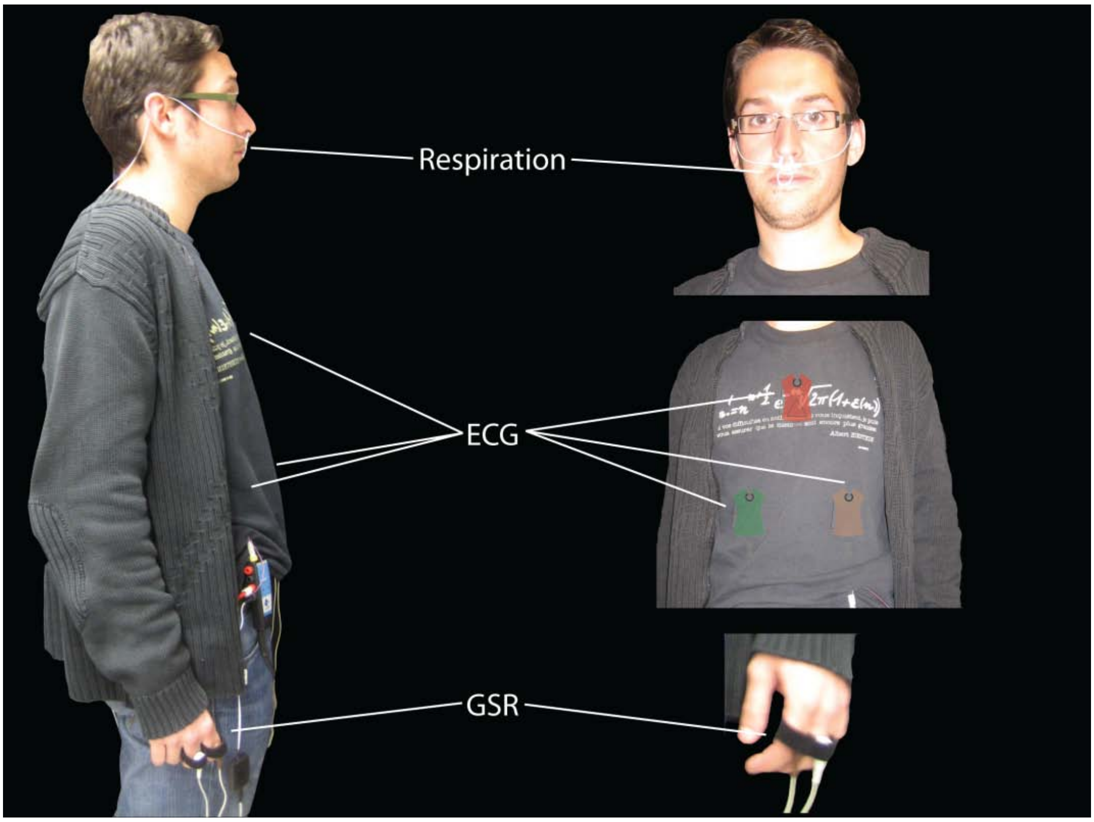

Figure 4-6. Front and detailed side view of a person wearing the three physiological sensors.

The data captured and streamed automatically to a Matlab/Simulink model via Bluetooth, in which all the relevant processing was carried out at the same rate as the incoming data (i.e. 256Hz). Logically, the Simulink model can be split into three distinct blocks: (i) data reception, (ii) individual filtering and analysis of HR, GSR and respiration, and (iii) finally data transmission to XVR. Each of these is covered separately in the remainder of this section. For each measurement we will outline how it is processed and quantized so it could be used to control bipolar events such as executing animations but also continuous adaptive colour changes. At every instance a vector containing all the current data that are being measured is received from the device via Bluetooth and initially split into n different scalars, in our case n = 3 for HR, GSR and respiration. Although initially we wrote our own ECG processing block to obtain HR and HRV, a proprietary block called g.HeartRate was made available by g.tec and it works in conjunction with the g.Mobilab+, which we used at a later development stage. It performs automatic filtering and processing of the incoming data stream and requires no user calibration.

In the next sub-sections we will describe each of the physiological measurements and how they were processed separately. In each of the sections we also sketch what portion of the data and how it was actually used to control the animations described earlier.

4.3.4.3.2 Galvanic Skin Response

Physiologically, GSR is a method for measuring the electrical resistance of the skin. Changes in the skin’s electrical properties are caused by events taking place in the environment and a person’s resulting psychological state. It can be measured from the human skin by applying a small but constant voltage to the skin [Dawson et al., 2000].

A recorded GSR signal is normally extremely noisy and requires extensive filtering prior to any analysis. We thus designed a digital infinite impulse response (IIR) filter of order 6 to remove noisy frequency bands. The filter’s response was similar to the response of a Butterworth filter and it is shown in

Figure 4-7. A sample sequence of filtered and unfiltered data is depicted in Figure 4-8.

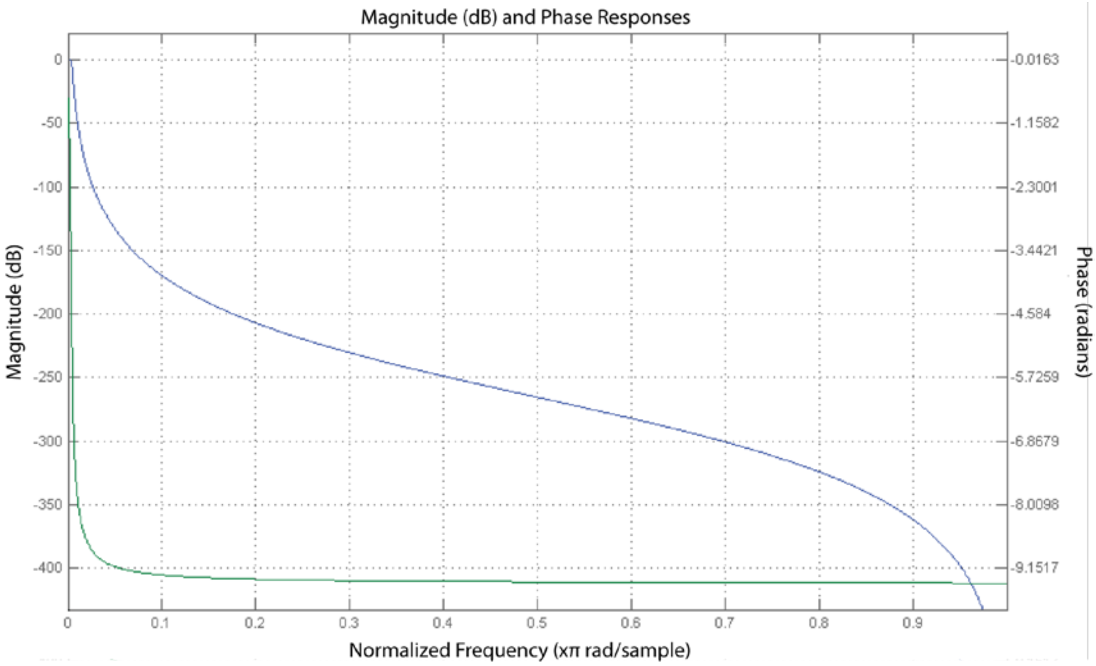

Figure 4-7. Magnitude and phase responses of IIR.

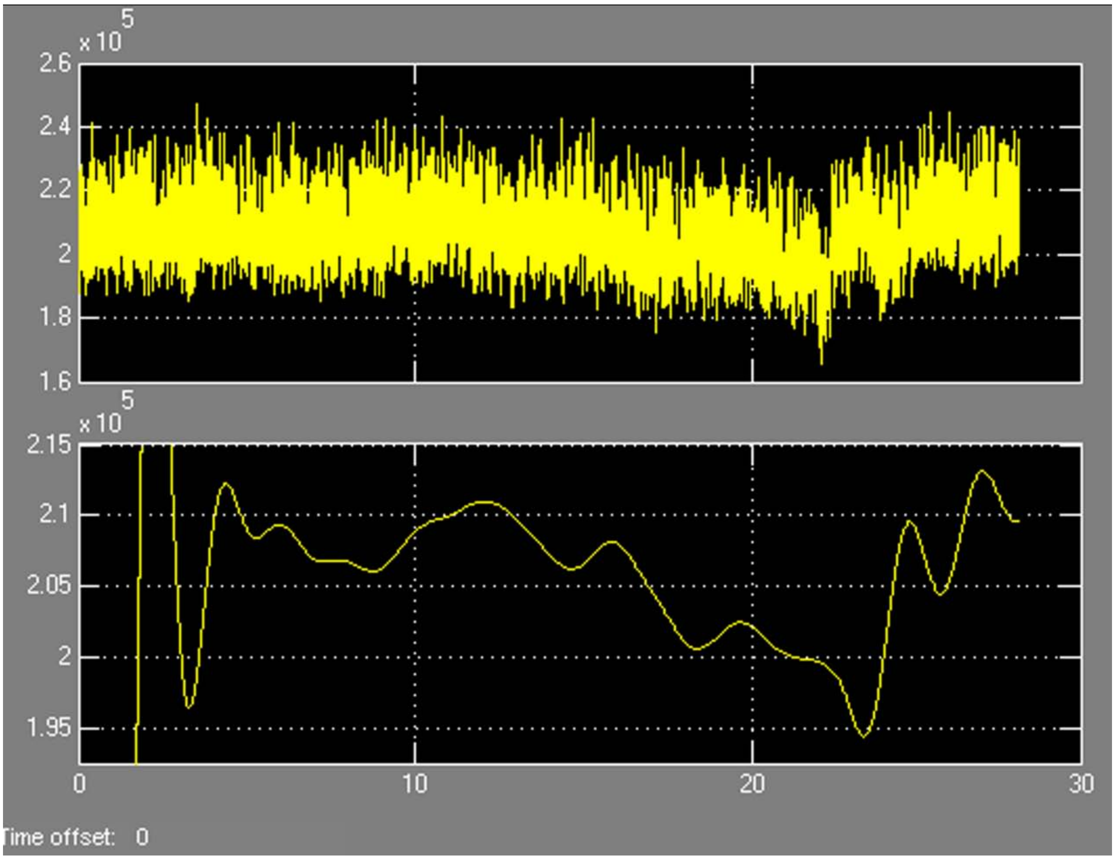

Figure 4-8. 30-second GSR raw (above) sequence filtered using an IIR filter (below).

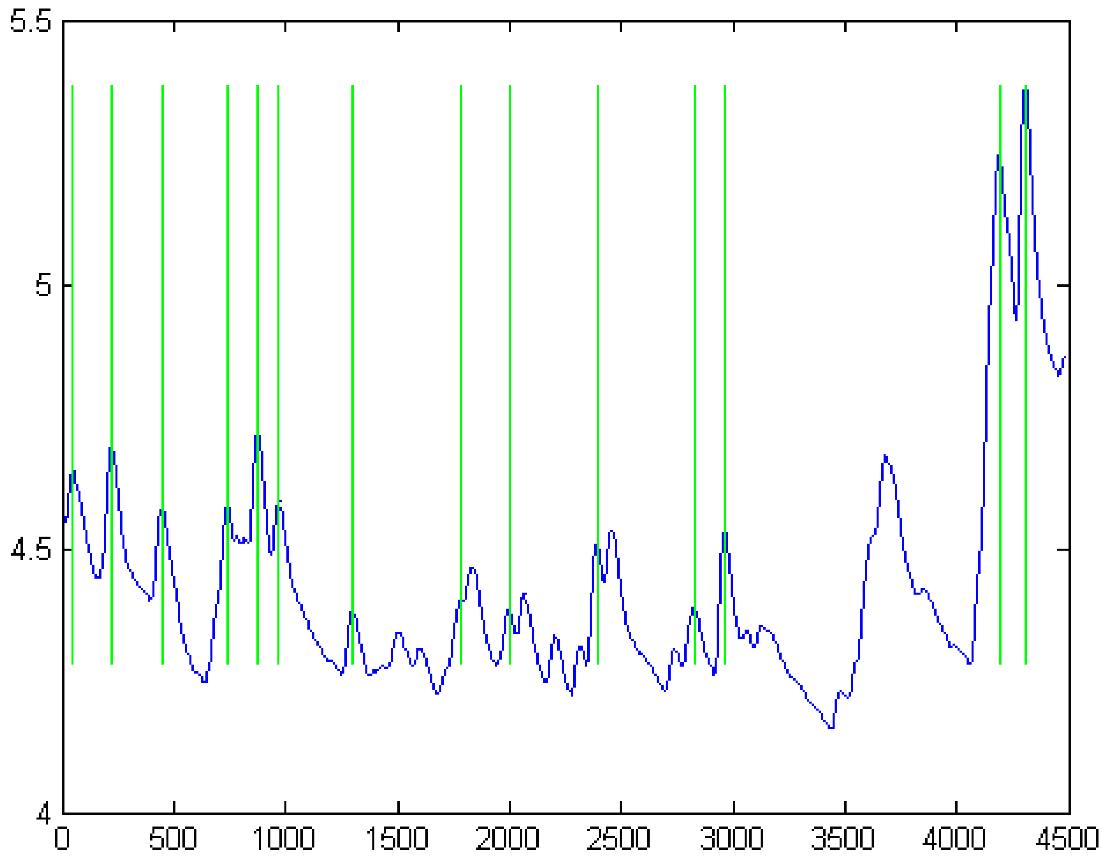

Figure 4-9. Filtered GSR signal of 4500 samples. The signal has been downsampled to 32Hz and therefore represents approximately 140 seconds of data. Green horizontal lines occur at detected peaks. Note that one obvious peak after around 3700 samples is not detected by the algorithm.

Next we analysed the GSR data in a windowed sequence of 15 seconds. Recall that the GSR can yield a measure of a person's level of arousal and in a given signal this is reflected in peaks [Venables and Christie, 1980, Nagai and Critchley, 2007]. See Figure 4-9 above for an example.

Given the sampling rate of 256Hz, the size of the window was 3840. We implemented a set of functions allowing us to detect the number and location of skin conductance responses (SCRs) in a windowed signal in real-time. It performs an additional layer of frequency filtering to remove additional noise. Since peaks are often located near the edge of a window we also add a 5-second margin on either side of the signal corresponding to the actual samples retrieved from the g.Mobilab+. Although some algorithms for detecting GSR response exist, for example using principal component analysis [Tarvainen et al., 2001], we chose to implement our own technique to find the peaks in the window. We first fit a spline to the signal and find its first and second derivatives using the Matlab functions csape and fnder. Now, we know from the second derivative which extrema are maxima and which ones are minima so we remove all the minima and are left with the peaks only. Of course, in a fairly degraded signal like the GSR it is not always trivial to find all the exact peaks, especially when this needs to be done in real-time, but nonetheless the algorithm performs reasonably well. As a rule of thumb, the shorter the signal the better the algorithm performed in finding the number and location of peaks (the latter however is not required and was only done in order to visualize the results). The detected peaks for a sample signal are shown in Figure 4-9 above.

The number of peaks per 15 or 20 second window are then used as a measure of arousal and the higher the number the greater the level of arousal. In an ordinary situation, a person would have 12 to 15 such peaks per minute. Under stress this number is likely to go up and this is often the case when people are exposed to new experiences such a VE. Now, our intention is to control a colour component of an avatar, which usually is in the range of [0.0, 1.0] and in order to scale the number of peaks to a value in this range we first perform a baseline reading in the beginning of each experiment and the average number of peaks per 20-second window marks the lower boundary (i.e. 0.0). Unless we seriously distress the participant in another pre-experiment baseline an individual upper bound is not as simple to determine however and we thus decided to average over the sum of detected peaks per 20-second window and this performed reasonably well.

5.4.3.3.3 Heart Rate and Foot Tapping

Heart rate relates to the number of heart beats per minute (bpm) and in a normal adult it is commonly between 60 and 100bpm in a relaxed state, though it can vary significantly based on factors such as gender, fitness or age. We use standard ECG measurement techniques whereby electrodes measuring an electrical impulse generated by the heart are placed on certain points on the skin. As stated above, we used proprietary software for generating heart rate analyses and we will not go into further detail here. The algorithm is capable of estimating HR and other figures in real-time and computations are based on a one-minute interval. Every heart beat is characterized by ventrical activity which results in a visible spike in the ECG signal called the QRS complex and the average number of such spikes per minute is essentially equal to one's heart rate. The time difference between two such QRS peaks can be used to describe the duration of a full foot tapping animation. An ECG sample is shown in Figure 4-10.

Since it refers to the time *between* two heart beats we designed a complete foot tapping animation to describe foot resting, foot lifting and foot lowering until foot rests on ground again (Figure 4-11). The terminal points, i.e. 0.0 in the beginning and 1.0 at the end of each animation cycle, occur at two consecutive heart beats. However, in this particular animation 0.0 and 1.0 are actually the same in terms of the bone rotations and positions, describing a complete animation cycle with two heart beats.

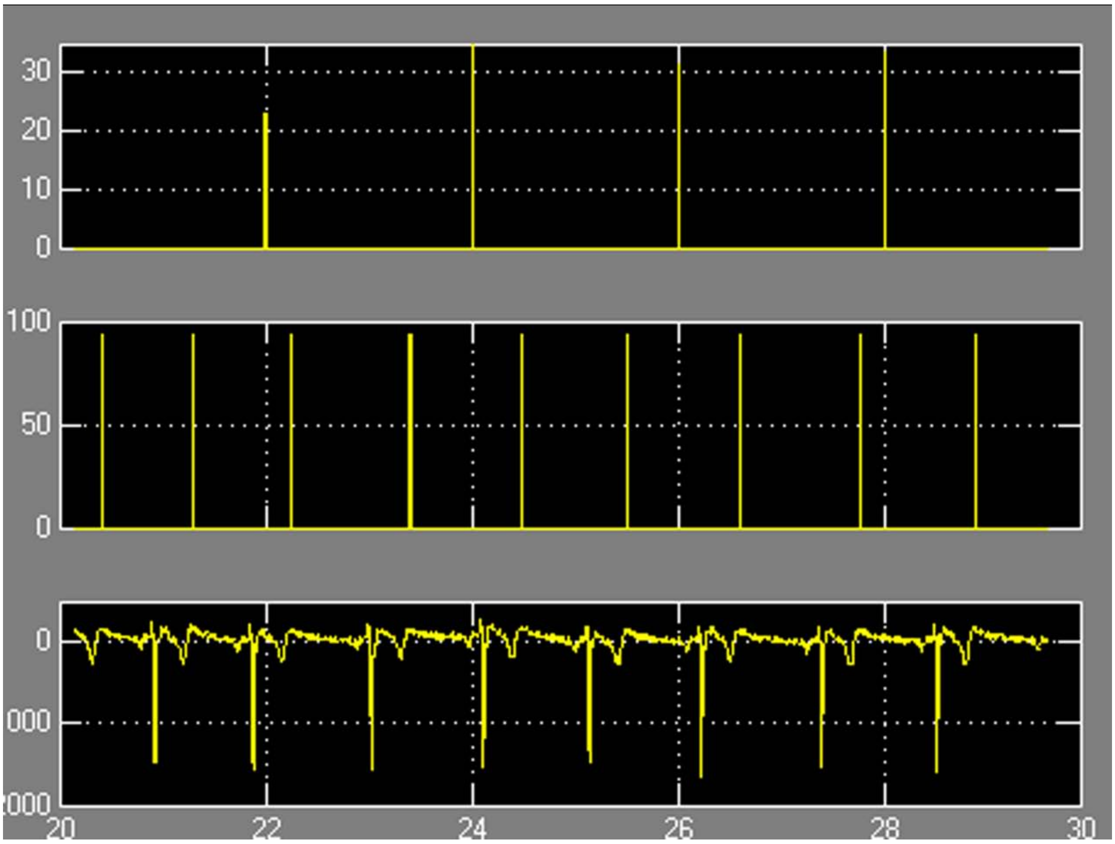

Figure 4-10. Ten-second interval of ECG data. The bottom graph shows the raw ECG data. The QRS complex occurs around the large peaks. The computed delay is shown in the middle graph and the topmost graph shows the calculated heart rate.

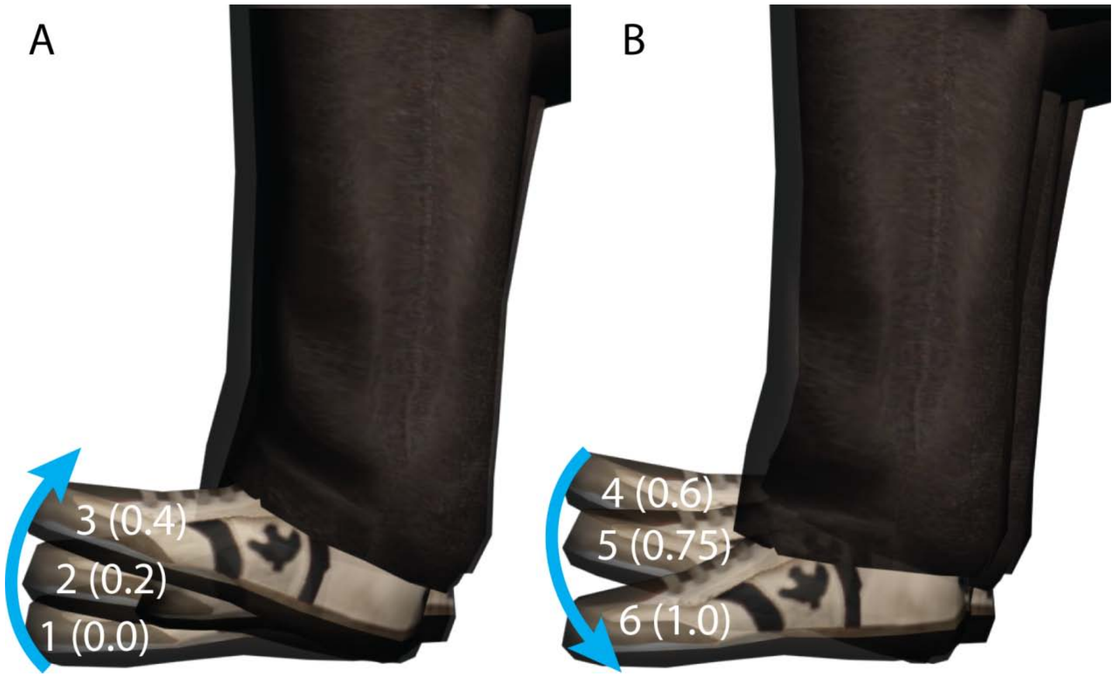

Figure 4-11. Stepwise time-compressed animation sequence of entire foot tapping movement. (A) shows upward movement for values ≤ 0.5 and (B) the downward sequence for values > 0.5. Note that (1) and (6) comprise the same bone configuration while (3) and (4) form a brace around the midpoint of the animation (i.e. 0.5).

Foot tapping is thus only computed once per heart beat by taking the time difference *d* between current and previous beat and dividing this by a pre-defined step size *s* (i.e. the number of steps the animation is broken down to) yielding a step timer *i* and the animation is incremented *s* times every *i* ms. Given the QRS difference in ms and a known and fairly constant frame rate, say 50fps, we can thus easily work out an increment by which the animation progresses at every frame, thus going from 0.0 to 1.0 in a finite number of steps of equal increment.

5.4.3.3.4 Respiration

Measuring respiration, the continuous intake of oxygen and outlet of CO_2 through the lungs, can be done in several ways. In our case, the g.FLOWsensor respiration sensor is originally intended to monitor the changes of temperature of breathing from nose and mouth. This allows us to infer a signal of approximately sinusoidal appearance, where minima and maxima refer to the exhaled and inhaled states of the lungs, respectively.

The incoming signal is filtered using a standard Butterworth filter and then we compute the first derivative of the filtered curve. The result is either zero, positive or negative. A peak occurs when the value is zero and we can infer whether it changes from negative to positive (i.e. a minimum) or vice versa (a maximum) by comparing its value with previous elements. If the value is non-zero there is no change in the derivative. Figure 4-12 shows how, using this method, we can fairly accurately determine the duration as well as beginning and end of an inhalation period.

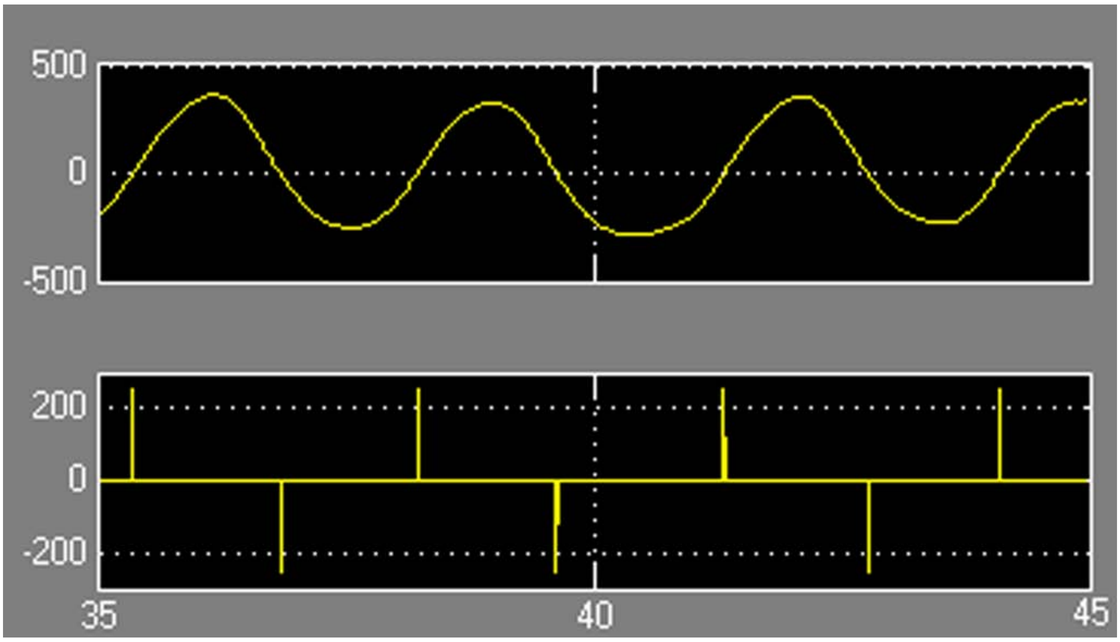

Figure 4-12. Ten-second interval of captured respiration signal. The top curve represents the filtered data, whilte the bottom one shows approximate locations of the zero crossings.

In order to model real-time respiration associated with a person's real breathing only requires us to know about the duration of one inhalation period. The procedure is similar to the one described in Section 5.4.3.3.3 above on connecting heart rate with a foot tapping sequence. In this example, however, we are not only interested in calculating time differences but rather our intention is to adjust for variations in the extrema that occur throughout the measurement process. One reason for this is that, unlike in the example on heart rate, the algorithm applied on the respiration signal does not always yield the exact location of the zero-crossing, so it is harder to synchronize between real and virtual breathing. Also, the signal should respond to the actual amount of breathing that a person does and not just the rate at which it happens. Going back to Figure 4-12, it can be seen that while the difference between each consecutive maximum and minimum, respectively, may not appear very large, small variations do amount to differences in breathing patterns. Some refer too more intake of oxygen and result in deeper and longer respiration than others and this is exactly why the respiration signal presented in Figure 4-12 above is not a perfect sinusoid but only an approximation of it. Calculating only the time step at which maxima and minima occur therefore does not fully capture the quality of the signal – heart rate in this sense is much simpler as we are really only interested in time intervals between two beats. If we played back the entire breathing animation for every minimum-maximum pair keeping track only of duration would show no major differences between deep and shallow breathing rhythms for example, so we need to come up with a different solution that takes into account also the depth of the breathing pattern.

Remember again that an animation can be performed using an arbitrary number of steps ranging from 0.0 to 1.0, where the former refers to the initial bone configuration at time t = 0 and the latter to the final configuration at the end of the animation, t = 1. The value 0.5 would result in the display of the midpoint of the animation and so on. Now, in addition to the duration of one respiration cycle we also keep track of a running average of maxima and minima which we then use as a basis to clamp the current values to animation endpoints between 0.0 and 1.0 (equal to the running average). The running average describes the long-term performance of the maxima and essentially smoothes the values. This way heavy breathing results in endpoints clamped to number very close the maximal morph targets 0.0 and 1.0 and almost the entire animation is played back within the given time frame. Shallow breathing, on the other hand, results in endpoints that are closer to the midpoint 0.5 and further away from 0.0 and 1.0, so that only a portion of the animation is actually played within given time constraints. The resulting animations appear much more realistic and closer to the real breathing pattern performed by the human participant.

4.3.4.4 Virtual Waiting Room

The model of the interior and exterior of the waiting room was prepared by using 3D Studio Max. Some of the components were freely available while others needed to be modelled. In addition, we had to apply adequate textures to most objects and light the scene to yield a realistic setup. Lighting the scene involves setting up and defining parameters of a number of suitable light sources. In a process called texture baking a new texture is generated from each texture in the scene that reflects the light sources and interactions with other objects such as shadows and reflections: all lighting conditions that affect a selected texture are taken into account and then it is rendered into a new texture. This yields static but realistic lighting conditions directly applied to the textures so no other actions have to be performed when the scene is rendered in XVR which does not offer interactive lighting. An example of texture baking is shown in Figure 4-13 (next page).

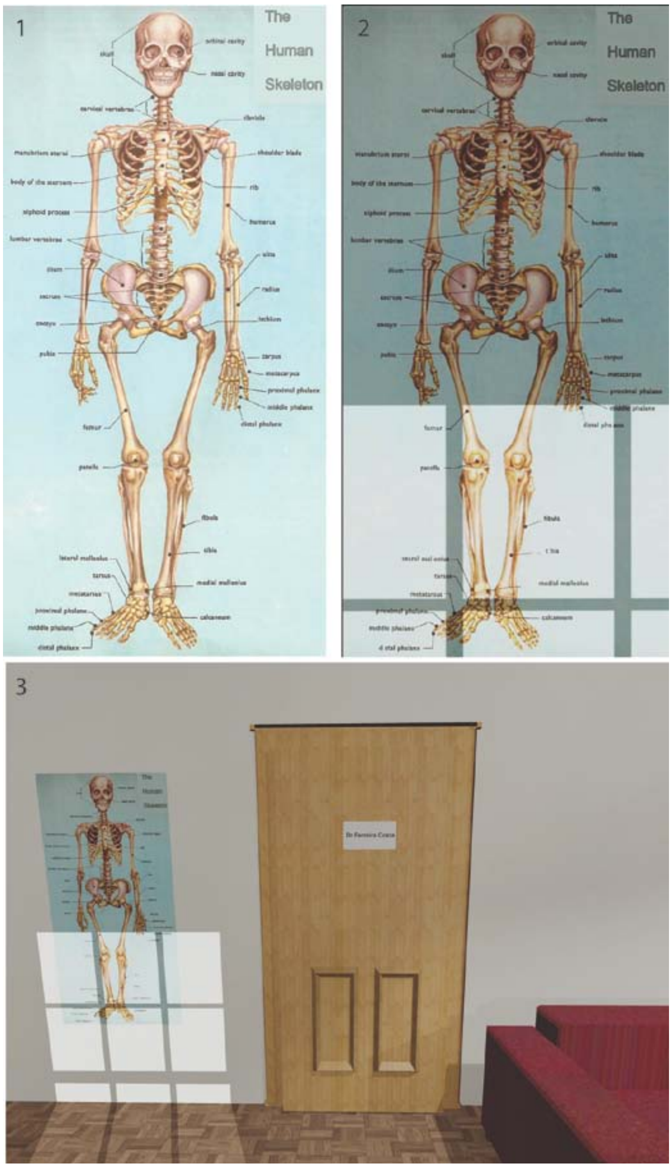

Figure 4-13. (1) Original texture of a skeleton poster, (2) baked texture with scene lighting and shadows from window frames (at the other side of the room), (3) scene render including the poster.

With the exception of the virtual characters and a revolving fan, every object in the model was thus static and not animated, so there was no need to export each mesh separately for use with XVR. A simple way therefore is to export the whole scene into a single AAM file, the native file format for XVR. While one file per scene object offers more flexibility in terms of real-time interactive animations, this was not a requirement in this environment. Exporting a scene into a single file can significantly increase rendering speed and given that we had ten virtual characters to animate interactively it made sense to reduce the cost of rendering the remaining scene.

The entire scene is enclosed by a sky dome and some simple outdoor geometry – cars, grass and pavement near the main windows. Clearly, the interior is more detailed than the exterior, and a basic bird's eye view of the interior is shown in Figure 4-14. The room is characterized by an entrance corridor and a main waiting area, which contains some furniture including sofas, tables, plants, paintings and lighting. The room was designed to be large enough to hold ten life-size virtual characters while also yielding enough space for a real human visitor. Figure 4-15 (page 120) offers an arbitrary perspective of the unpopulated room while Figure 4-16 (page 120) offers another view of the room populated with the ten characters.

We made extensive use of this model during pilots (Section 4.3.3) but some volunteers commented that some of its details influenced their decision. We thus created a simple environment which we used for the experiments. The environment was built in a similar way to the method described in this section and an image can be found in Figure 4-17 (page 121).

Figure 4-14. Bird's eye overview of the waiting room.

Figure 4-15. Perspective from one corner of waiting room.

Figure 4-16. Final scene, including waiting room model and populated by ten interactively animated virtual characters.

Figure 4-17. Image of the simple environment used in the experiments. It does not comprise any furniture or windows except for three sofas placed in a semi-circle.

4.3.4.5 Networking Protocol and System Architecture

Recall our remarks in Section 4.3.4.1 on data transfer and that we initially used a different device for physiological monitoring (i.e. Nexus-4). While the device is shipped with its own software and a DLL for custom program integration, the use of XVR for rendering and Matlab/Simulink for data processing made it somewhat difficult to directly exploit the DLL functions, especially since the raw data was first going to be processed in a Simulink block which then passed on some vector to the VR engine.

The main motivation for using a networking protocol is that there are three separate modules involved which each run on separate machines. The capturing device is independent from the machine handling the data processing and there is a second computer dedicated for rendering the VE. The last two topics were covered in Sections 4.3.4.3 and 4.3.4.4, respectively, and in this section we will see how these two and the third module (i.e. the capturing device) interact with each other. Now, each of the three modules needs the ability to pass on data to the next process in line. Figure 4-18 outlines the basic program flow and illustrates the need for a suitable networking architecture. The Physio-PC thus served as a client machine to the incoming data and a server hosting the data transfer from itself to the Rendering-PC.

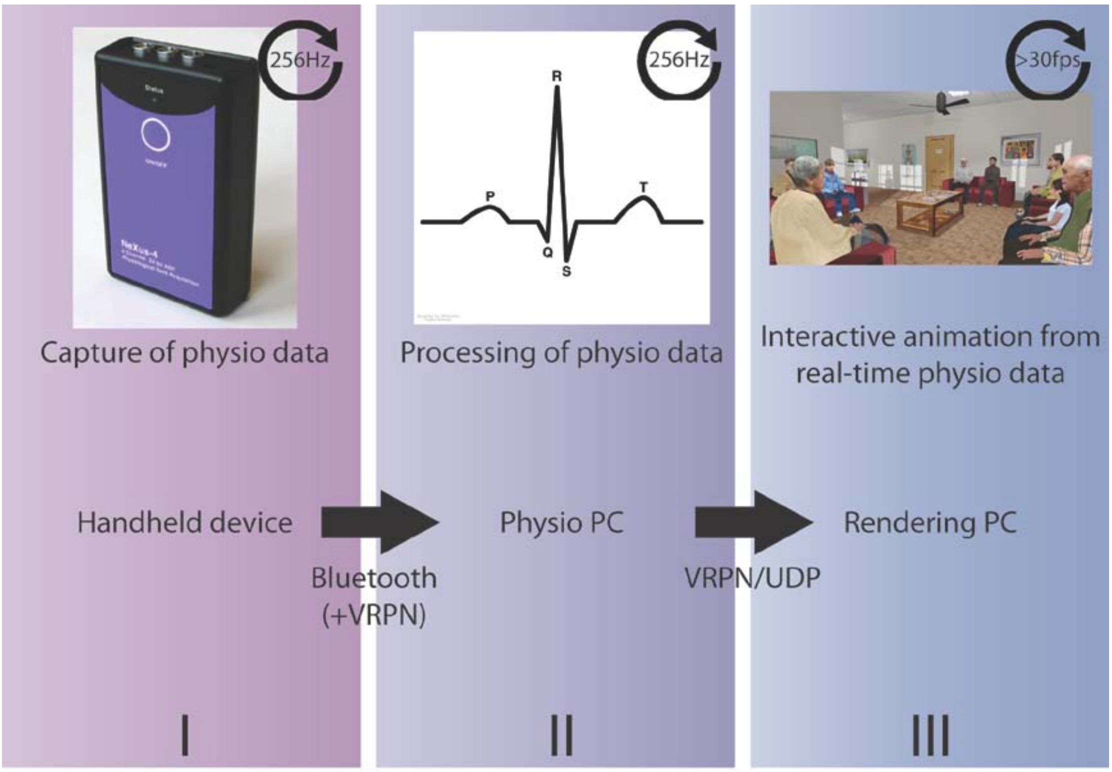

Figure 4-18. Simple diagram of data flow and networking requirements. The left-hand block (I) is a device for capturing physiological data, which is transferred to a PC (II) via Bluetooth, where the data is processed. The resulting feature vector is then passed to another PC (III) where it is used to animate objects and the entire scene is rendered. Each block also shows the sampling or frame rate at which it operates.

In previous experiments we successfully used the Nexus-4 device and the DLL functions in conjunction with the VRPN library [Taylor et al., 2001] to access real-time data from within XVR, so we implemented a similar version that would allow us to access readings from within Simulink and also pass on our results to XVR. Now, the first problem we encountered was that simulation time in Simulink is most likely not equal to time in seconds or milliseconds and it rather advances at an arbitrary rate, essentially as fast as it can. This is due to the nature of Simulink being a platform for creating simulations and prototypes and it is not very common to rely on it for impromptu data processing. While we solved this problem by manually and explicitly discretizing simulation and sample time thereby enforcing this through a dedicated custom-built function block, it required substantial changes to be made to our existing software architecture.

Embedding the DLL that was provided for the Nexus-4 into Matlab did not work at all due to software conflicts, so we decided to bypass it and retrieve the data from a server created from the VRPN library, which in effect created another client/server application run on the same machine (i.e. the Physio-PC) – hence the remark on VRPN between I and II in Figure 4-18. However, even this approach failed and the whole Simulink model became unstable as a consequence of this.

It became rapidly clear that the Nexus-4 device and Matlab/Simulink were not be able to exchange data at the required sampling rate without investing a serious amount of extra effort, and therefore we decided to use a different device. At this point we could have equally implemented the software dealing with physiological processing on a different platform such as C++ and continued using the same device but we finally decided to implement the former option. We thus replaced the Nexus-4 device with a g.Mobilab+. It has the advantage that it is already shipped with a Simulink block which can be readily used as the basis for an arbitrary Simulink model. Also, the device itself directly controls the sampling rate and governs the speed at which data is received and processed in Simulink, which is a constant 256Hz. It also computes HR, HRV and do further statistical analysis on the incoming ECG data.

Now, data was transfer from the device to the Physio-PC was performed automatically once the device was connected and we only needed to focus on providing a method to transfer data from the Physio-PC to the Rendering-PC. We first tested our existing method using the VRPN architecture but it failed to deliver the data at a constant rate. We thus replaced VRPN with a simple UDP protocol which allowed us to send character strings over a network without delay.

Now that we have covered all aspects of the software implementation we will briefly sketch the program flow and communication. The software modules, as depicted in Figure 4-18 above, suggest a unidirectional data flow starting from the device transmitting readings to the Physio-PC, which in turn passes on its results to the Rendering-PC. Since the last module runs at a much slower rate than the other two, i.e. around 30fps compared to 256Hz, the Rendering-PC receives data whenever it can which is not necessarily at a constant rate, since its speed depends on the rendering functions. However, this operation is fairly safe considering that there are no sudden or unpredictable changes in HR, GSR or the respiration signal. Also, as we saw in Section 4.3.4.3 the values of a feature vector do not change at the same rate as the sampling rate would suggest, and passing on this vector to the XVR environment in fact happens at a much slower rate because each result depends on a *n*-second windowed operation.

4.3.5 Task

After fully equipping them with the physiological sensors and the device, participants were given the task to enter the virtual waiting room and observe the *n* virtual characters as long as they liked and to identify the avatar that was controlled by another person:

> *"You are going to be seated in a waiting room. You will see 6 people there. None of them can see you and they are not aware of your presence. One of the characters is a real person who controls the character from somewhere else. That person also does not know they are being observed. Look at the people and identify the character that is controlled by the real person. Unless you are sure about the character beforehand, at the end of 10 minutes we will stop you and ask you. This is an investigation and not in any way a test of you."*

As explained above there was no other person in control and it was simply their own behaviour that determined (parts of) the environment. We therefore had to state a reason for monitoring their physiology so that it would not become obvious why we were doing this.

4.3.6 Population

A total of 13 participants (three males and ten females) were recruited for this experiment. All participants had normal or corrected-normal vision and the study lasted approximately 45 minutes.

4.3.7 Procedure

The pre-experiment stage consisted of three objectives. The first was to familiarise the participant with the equipment, in particular the tools for physiological measurements but also the powerwall and the correct use of the wand and passive stereo glasses. Participants then signed a consent form stating that they had been given enough information about the study and that they agreed to take part in the study. They also filled in a questionnaire on demographic information. This form is printed in Appendix D.

After this, a voice recording was played back to them attempting to produce a higher state of relaxation and awareness.

In the experimental phase the participant is seated in front of the powerwall. He then enters the environment and attempts to complete the task read out to them (cf. Section 4.3.5). One of the characters appears to be highlighted while all the others seem to be slightly shaded. This is for concentration purposes only and volunteers can choose which avatar they want to highlight by pressing a button on the wand. In addition, we asked the each participant to speak to the currently highlighted avatar even though they did not reply to them. We introduced this measure because GSR changes during speech are greater than during silence. This had the effect that facial blushing became more noticeable during speech.

Even though they were told that they could stop at any time, the experimenter remained silent in the background during this phase and ignored the participant’s comments unless he really was sure he had identified the correct avatar or wished to abandon the experiment. None of the experiments exceeded more than ten minutes.

After the experiment finished the participant was asked to identify the correct avatar. The result was noted down and there was also a five-piece questionnaire on presence participants filled out before being debriefed and paid €5 for their participation.

4.4 Results

None of the 13 participants selected the R-avatar and instead chose one of the dummy avatars. The character that received the most “votes” was a black woman (five decisions). As already established during the pilots and outlined in Section 4.3.3, a large number of people seem to be swayed towards elderly people and six people chose one of the elderly avatars. Although the majority of participants chose a character that was seated on either the left or the right sofa (ten decisions) the sample size was too small to draw any conclusions from this data. Presence scores are presented in Table 4-3 below.

Table 4-3. Mean scores and standard deviations for the five questions on presence.

Question	Mean	SD
Q1: I had a sense of being in the virtual room (1=not at all ... 7 all the time)	4.5	1.6
Q2: To what extent were there times during the experience when the virtual room was the reality for you (1=never ... 7=all the time)	4.6	1.6
Q3: When you think back about your experience, do you think of the virtual room more as images that you saw, or more as somewhere that you visited? (1=images ... 7=somewhere I visited)	3.7	1.4
Q4: During the time of the experience, which was strongest on the whole, your sense of being in the virtual room, or of being in the real world of the laboratory (1=laboratory ... 7 = virtual room)	4.0	1.1
Q5: During the time of the experience were you more aware of being in a laboratory or being in the virtual room? (1=laboratory ... 7=virtual room)	4.5	1.5

4.5 Discussion

As already mentioned above, the results from the pilots clearly contradicted our initial hypothesis. Evidently, the task of identifying the real part in a VE was too difficult, which poses several questions. The same is true for our actual study which was in some ways simplified compared to the pilots.

There are several indications for why the results turned out this way. First of all, it is likely that the task given to the volunteers was either too difficult, too vague or too misleading and that they were looking for the wrong signals. Also the motion of all avatars was determined by real, either pre-recorded or live, physiology which could lead to the somehow valid conclusion that all avatars presented a similar degree of realism – only one however referred back to the actual participant. On the one hand, it could be partially due to the task given to the participants (cf. Section 4.3.5) and that a real part was in fact present in every avatar although only one of them reflected on the participant's physiology in real-time. So the task could be either misleading or not clear and precise enough. One the other hand, part of the ambiguity may have been introduced by the fact that, while participants were actively scanning the VE for clues as to the identity of the "real" avatar, the task is essentially intuitive and introspective in nature and requires a great deal of (at least subconscious) inward attention. According to the task it is simply a matter of taking the right conclusions about some subjective observations made in the VE, and while this is true, it largely obscures the fact that, were the stated task true and one avatar was really controlled by another person while all others remained automated, we speculate that the results would very likely be the same as the present study. Although it has to be stated in an abstract way the task becomes misleading. To a degree, this view is supported by the fact that none of the participants ever questioned the use of physiological measurements.

Second, the animations at times appeared slightly artificial. This was mainly a problem with the breathing animation and not so much regarding foot tapping or even facial colour. Due to hardware constraints there was a small but noticeable latency regarding the respiration signal and the relatively low agreement between real and animated breathing when not breathing deeply could be the result of a combination of this, the algorithm estimating the current respirational cycle (cf. Section 5.4.3.3.4) and the realism of the animation itself (cf. Section 4.3.4.2). However, deep and regular breathing resulted in good and noticeable synchronization of animated breathing. Regarding all three correlations it is also possible that they were not noticeable enough and thus may require a good deal of exaggeration[18] so that individual correlations become more apparent.

Third, it is possible that the combination or the types of physiological measurements we chose were simply not adequate given our goal. Perhaps we should begin experimentation with a smaller number of measurements, possibly only one, and verify that they the mapping leads to the desired outcome.

Given the relatively high agreement on the choice of the R-Avatar – throughout the pilots and the experiment the majority chose one of the elderly avatars – it is possible that participants who hesitated to make their choice automatically selected one of the apparently weaker characters. Also, roughly half the participants in the study selected the only black avatar. A future study might take this into account by collecting data on personality traits such as the Five Factor Model [Costa and McCrae, 1992] with which we may be able to correlate with the volunteer's choice.

4.6 Chapter Summary

We introduced a VE that enables its visitors to interactively induce changes to the environment via their own physiology. The VE contains a number of virtual characters which can either respond interactively to the participant's physiological states as determined by his heart rate, respiration and GSR, or an arbitrary combination of them. We developed and employed several algorithms for real-time physiological processing and analysis in order to use these data to control pre-built animations that were looped throughout the experience and scaled in various ways, affected only by the participant's physiology.

In a between-groups study with eight and five subjects, respectively, we found that it was not possible for the vast majority of the participants to identify their

[18] Special effects and games industries often face the problem of a lack of realism in some of their scenes and applications despite the use of accurate real-world physics. Similarly, the uncanny valley hypothesizes that robots or avatars that almost look and act like real humans will be more easily rejected by humans and often cause a strong negative response [MacDorman, 2005].

own physiology or its virtual manifestation. Possible reasons for this were discussed in detail in the previous section ranging from problem complexity, task statement over to technical issues and distraction. The two main problems may have been task statement and distraction.

This currently implies that we cannot employ physiological measurements in order to design a more realistic VR experience or one that is more tailored to individual responses or that can assist in interactively guiding humans through a scenario.

Regarding future work, there are a lot of things that can be done to enhance this project. A different experimental design may have brought more clarity to the subject. For example, testing the general possibility of implicit physiological interaction could be carried out in a much more straightforward environment with geometric or otherwise simple objects that respond with simple changes (e.g. colour, scale, rotation) to a person's physiology. Once a connection has been shown and established, a more intricate set of responses within a more realistic environment could be designed and tested. One could also present only one stimulus (in our example only one avatar) at a time when evaluating its realism and be able to switch from one to the next. Such experiments could give rise to further techniques that can gradually transcend into more complex guidelines including elements of storytelling.

5 Mental Interaction – The Smart Home

5.1 Synopsis

In this chapter we present a non-invasive brain-computer interface (BCI) for controlling a VE, in which the human user learns to focus his thoughts and interact with the VE solely through mental processes. Unlike the project described in the previous chapters, the participant does not use any other (physical) aspect or part of his body except thought in order to communicate intent to the environment. This work was developed in collaboration with g.tec OEG[19], one of the partners of the PRESENCCIA project.

The main idea behind this project was to evaluate the use of a wireless BCI system as a global interface for VEs. As we mentioned already in Section 2.6 due to its slow information transfer rates BCI research is mainly directed at rehabilitation and used as an assistive device for severely impaired humans. Much research is undertaken with primates [Lebedev and Nicolelis, 2006] but also in human beings [Wolpaw et al., 2002]. BCI systems are used mainly for moving a cursor on a computer screen and controlling external devices or for spelling purposes [Krusienski et al., 2006, Guger et al., 2001, Vaughan et al., 1996]. Our system is based on the P300 potential [Sutton et al., 1965]. We already provided some background information on the P300 and other BCI techniques in Sections 2.6.3 and 2.6.4, respectively.

This research encompasses two studies that each deal with a separate aspect of usability of BCIs in VR. Our studies focus on a virtual apartment, a so-called smart home, where a number of appliances can be automated and controlled from a remote location. One study plainly assesses user performance and overall usability in the smart home environment while the other evaluates how its use affects the sense of presence. The latter study is carried out by comparing the method with a traditional means of interaction. In the remainder of this chapter we will first discuss our motivations for carrying out studies followed by an in-depth examination of the experimental setup and

[19] http://www.gtec.at

procedures for both experiments. We first summarise the usability study in Section 5.3 and the presence study is outlined in the following Section 5.4.

The work received ethical approval from the Ethics Committee of Clinical Research in Barcelona (Comité Ètico de Investigaciòn Clinica – CEIC) and has been published [Guger et al., 2009, Edlinger et al., 2009, Guger et al., 2008, Edlinger et al., 2008] and submitted for publication [Groenegress et al., 2009a].

5.2 Aims and Expectations

As we mentioned earlier (cf. Section 2.6.5) Bayliss and colleagues already presented a BCI interface for a smart home environment [Bayliss, 2003, Bayliss et al., 2004]. The work however only acted as a proof of concept demonstrating the technological feasibility of such an installation by comparing its use within different systems: inside all-enclosing HMD or viewed on a monitor. The work therefore does not directly deal with usability and user performance in a pure VR setup but rather compares between an immersive and a non-immersive one, which clearly offers no insight about its viability as an interaction device for VR. Another coupling of BCI and smart homes was demonstrated in [Leeb et al., 2007a]. However, motor imagery was used instead of the P300 interface.

Our work differs from earlier approaches by studying in detail usability and user performance. In a second study we also evaluate how the use of a P300-based BCI affects one's sense of presence by comparing subjective presence measures from participants who control the environment using a BCI versus participants who use a wand and gaze-based methods to manipulate objects. The two interactions were designed to be similar to each other and the actions based on gaze are similar to the P300 approach. Instead of looking at symbols representing actions on a screen gaze is used to select virtual objects directly from the VE. Interactive objects implicitly convey an action, e.g. an open door can be closed and vice versa. Both studies are guided as opposed to self-paced and the participant is given a fixed sequence of tasks.

There are several reasons for this two-part study. First of all, one motivation stems from the fact that to our knowledge no comprehensive study on usability combining VR and BCI use has been performed to date and the reasons for this are obvious: Both fields are still in a relatively early phase of development and BCIs have essentially not advanced to an extent that permits their use beyond the laboratory because they are fairly complex to set up, maintain and they are also prone to errors. Secondly, present BCIs are slow compared to most other interaction devices and the information transfer rates are usually much less than 100 bits/s so there often is no significant gain in using them compared to other devices, at least not when the user is healthy. This is precisely why BCI research and development has found a niche in the rehabilitation sciences but not much elsewhere.

Another potential benefit for building such an application in VR is rapid prototyping. It is a lot more flexible and cost-effective to design and test a complete system in VR first before implementing the actual technology in a real setting.

Finally, we aimed to demonstrate that such a system can be used by a naïve user with very little training. This is an important point at least from the perspective of a BCI experiment, where training often takes weeks or months – our studies were performed on the same day after a one-hour training session.

5.3 Usability and Performance Study

5.3.1 Variables

We conducted a within-group study in three conditions where each participant faced each condition in the same order. The difference in the conditions was the number of flashes used to classify a symbol and subsequent action: 8, 4 and 2, respectively. For each condition we measured performance in two ways: First we recorded the average performance per participant over all tasks and matrices. Remember from Section 2.6.3 that a P300 matrix consists of a grid of symbols that each flash randomly at certain time intervals. The basic speller matrix is printed again in Figure 5-1 below.

Figure 5-1. Original P300 speller matrix with symbols, in this case alphanumeric characters, aligned on a grid.

Second, we also recorded matrix performance that is the number of true positives in each matrix for all participants. This way we were able to calculate average performance levels for all participants as well as individual performance, but it also allowed us to highlight which matrix yielded good or bad performance. This is an important marker since we used non-uniform matrices instead of aligning symbols along a rectangular grid and thus wanted to be able to track how well each new configuration fared compared to the others. Finally, the three conditions are designed to highlight that despite short training periods satisfactory results can be achieved for a different number of flashes.

5.3.2 Piloting

In previous experiments we piloted the study in several stages in a CAVE-like environment and on a powerwall. In the first part we were mainly concerned with the viability of the technical setup and the main goal was to verify that the filters of the stereo glasses did not interfere with the correct perception of the P300 display and its flashing symbols. If latency or other errors would have been introduced by this method we would have had to find alternative methods of presenting the VE.

Feedback from participants convinced us to carry out the experiments in front of a powerwall instead of using the CAVE as the latter setup required us to place a table for the laptop displaying the P300 and chair for the participant to sit on into the CAVE, which severely distracted them from experiencing the VE that surrounded them. The CAVE setup included a small table and chair placed in the middle of the CAVE. The laptop displaying the P300 was placed on the table and the participant would sit down in front of the table. The difference to the powerwall was that they were seated and that the P300 display was in front of them. Although the display could have been placed anywhere inside the CAVE, it was necessary to place it on a table and since the participant was not allowed to move much he or she was seated in front of the display.

During the first set of pilots the training methods were different from those employed during the eventual study and a single classifier was computed from several spelling tasks meaning that the layout of the training matrix was different from the different matrices employed in the experiment (see Section 5.3.3.2 for more detail). Since our environment allowed for more than 200 operations, we used several matrices instead of just one. However, during training we experienced several problems and therefore decided to compute a classifier for each of the experimental matrices used in the study. This substantially extended the duration of the training period but eventually resulted in a more robust set of classifiers.

5.3.3 Apparatus

5.3.3.1 Equipment

We used a g.EEGcap to mount eight electrodes to the participant's head. The electrodes, in turn, were attached to a g.MOBIlab+ for biosignal acquisition and wireless Bluetooth transmission[20]. The g.MOBIlab+ is a small device that can be carried around the belt, allowing its wearer to move around freely in the laboratory, see Figure 5-2 for an illustration of the two devices.

A proprietary Matlab/Simulink[21] model was used for acquisition, analysis and classification of the EEG data. The algorithm essentially detects the most likely P300 response during each iteration and associates it with the signal highlighted 300ms before. The candidate responses are accumulated and evaluated at the end of each cycle. There should be one candidate per iteration and the operation with the highest number of candidates is selected and a decision is formed.

Figure 5-2. Image of the g.EEGcap and g.MOBIlab+.

[20] http://www.gtec.at
[21] http://mathworks.com

The P300 interface is displayed on a separate computer screen, a handheld device, a stationary PC or a laptop and throughout the experiments we used a laptop monitor. The VE is displayed on a 3x2m powerwall and the human head is tracked via a six degree of freedom (6DoF) Intersense IS900 motion tracker[22], attached to a pair of passive stereo glasses that are worn by the participant in order to perceive the scene in full 3D. Also, it was important that the glasses did not impede the perception of the P300 flash cycles displayed on the other screen and this was tested during trials.

5.3.3.2 Implementation

We designed seven distinct P300 matrices that overall have a less constrained layout than the original speller interface. Although there was no evidence that the P300 interface can only work in conjunction with a symmetric map. This had never been tested before. The main reason for this is that the P300 has almost exclusively been used for spelling before. Also the layout and number of icons (i.e. possible interactions) varied for each interface: being 25, 50, 30, 38, 40, 13, 22 symbols, respectively. As outlined in the previous section, the icons in the currently displayed interface would randomly flash for a short time period. By concentrating on a single icon, a measurable response is generated in the brain whenever this happens. An example of an interface used in this study is shown in Figure 5-3 below and the complete set is presented in Appendix D. They are music, movement, lighting, goto, temperature, TV and phone. Some of them bear some similarity (e.g. music and lighting) while others vary greatly (goto and movement). Our intention here was to keep as much overlap and key functionality between different matrices if possible and to logically separate navigation from other operations.

[22] http://isense.com

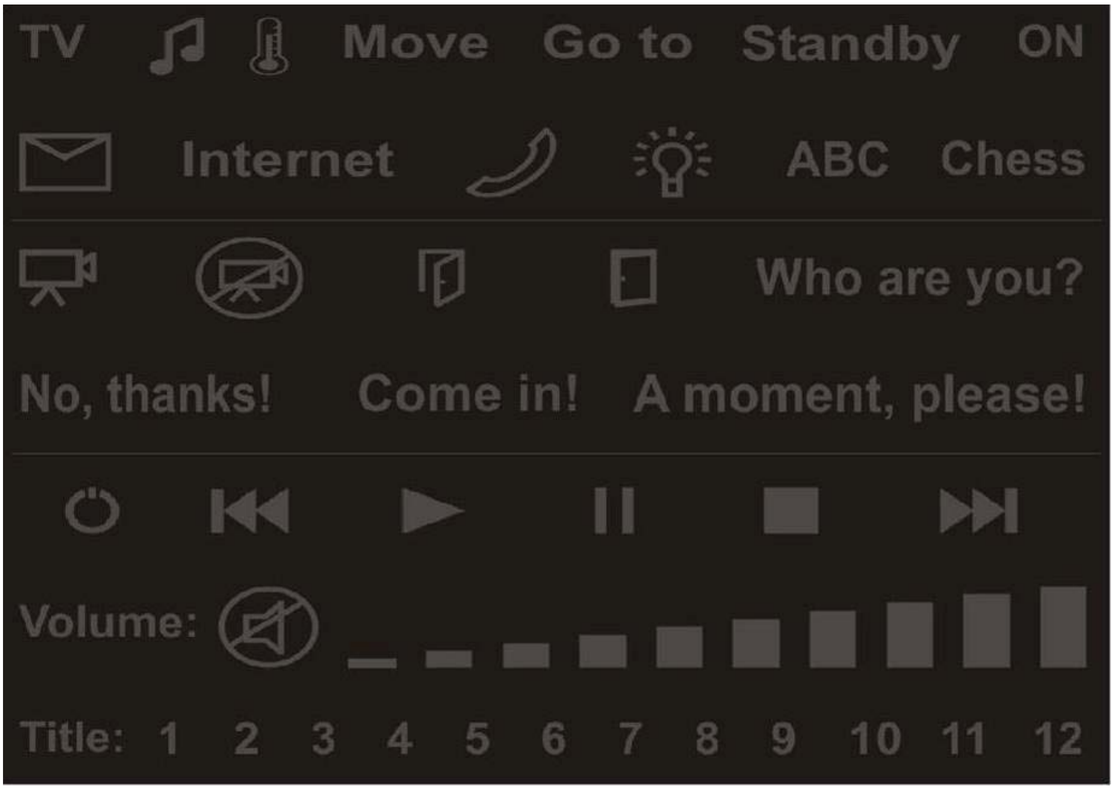

Figure 5-3. Music matrix, one of the seven P300 Matrices used in the experiment. It offers 50 symbols as a mix of icons and words denoting possible actions that can be executed. Note that not all of the symbols are implemented as actions in the VE.

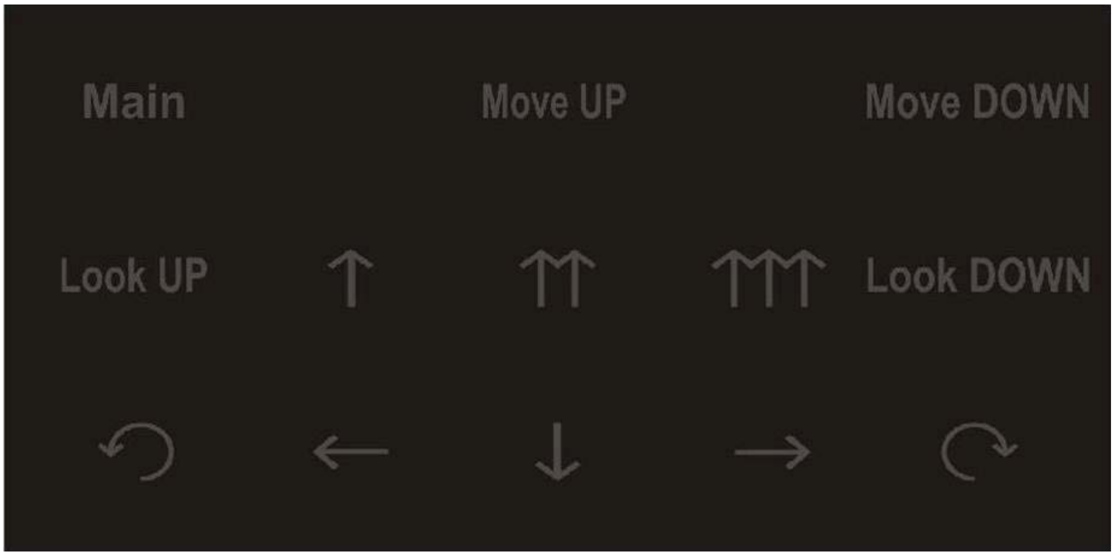

Figure 5-4. Movement matrix with 13 symbols representing common navigational operations.

XVR was used to control certain objects including lighting. The interactive objects were: four doors, two lamps, a television set, a stereo system, a telephone and an air conditioning system. Most of these were implemented with real-life resemblance, so the stereo system would account for standard song and volume operations to be carried out (e.g. play, pause, stop, forward, previous, volume up and down), and every operation was conveyed in at least a visual or audiovisual cue. For example, making a telephone call would display the dialled number onscreen, different temperatures of the air-conditioning system were represented by a light indicator on the system, switching channels on the TV would result in different videos each representing a different network being shown in the TV frame and so on. Table 5-1 (next page) shows a comprehensive list of operations and their mappings.

Some interactions required less attention than others. For example, while operating an air-conditioning unit, a stereo system or a television set are roughly the same in terms of possible choices (e.g. on, off, change programme/temperature/song/volume), others were either binary (e.g. open/close doors) or unary/open-ended (i.e. navigation). Regarding navigation we implemented slow and rapid movements in four directions plus turning left or right by 90 degrees. These actions were incorporated into a single Movement matrix shown in Figure 5-4 (previous page). Since movements resulting from the Movement matrix were fairly slow to execute and repeat, we implemented another matrix allowing participants to teleport to one of 21 different landmark locations inside the apartment, thus accounting for small- and large-scale movements.

Once a decision is made it is passed on to the VE and the operation is executed immediately. For some matrices not all of its listed symbols/actions appearing were implemented in the virtual apartment. A list of all symbols that are used in the VE (or for changing matrices) is summarised in Table 5-2 (page 141).

Table 5-1. Mapping between selected action and audiovisual response of the environment. Most relations, or at least the audible and visible response of the environment, are self-evident.

Category	Visualization	Type
TV Channel (1-7)	Displays different videos on television set depending on channel selected, includes audio	Visual, Audio
Track (1-12)	Selects and plays selected audio track	Audio
Volume (0-10)	Audio volume of the selected device (TV or stereo) increases or decreases	Audio
Temperature (18-25)	A light panel on the air-conditioning unit indicates the current selected temperature.	Visual
Door (3 doors)	Animates Opening or closing of selected door accompanied by audio.	Visual, Audio
Phone Call (9 numbers)	Onscreen display of person and/or number called with audible dial tone.	Visual, Audio
Teleport (21 locations)	Raises viewpoint to some distance above the apartment, so that whole apartment covers field of view, then moves towards selected landmark eventually settling there.	Visual
Light (1 lamp)	Displays different textures reflecting the current lighting parameters, according to whether lights are switched on or off. Accompanied by sound resembling light switch	Visual, Audio

.

Table 5-2. Possible actions that can be effected through each of the matrices and number of total operations that are reflected in the VE. Note that some matrices contain symbols that are not implemented in the apartment.

Matrix Name	Possible Actions
TV (36 operations)	Select another matrix[23] , Common operations[24], Power, Channel (1-7), Volume (0-10)
Music (46 operations)	Select another matrix, Common operations, Power, Previous, Play, Pause, Stop, Next, Volume (0-10), Title (1-12)
Temperature (30 operations)	Select another matrix, Common operations, On, Off, Temperature (18-25ºC), Bathroom door, Bedroom door, patio door
Phone (26 operations)	Select another matrix, Common operations, Ambulance, Fire, Police, Caretaker, Mother, Father, Wife, Ex-Wife, Girlfriend
Light (18 operations)	Select another matrix, Common operations, Living Room
Movement (13 operations)	Main[25] , Move Up, Move Down, Look Up, Look Down, Forward 1, Forward 2, Forward 3 steps, Left, Right, Back, Turn Left, Turn Right
GoTo (22 operations)	Main, teleports to 21 different landmark positions throughout the apartment from A-U

Table 5-1 sums up how actions are illustrated by audiovisual effects. For most of them the binding is straightforward and very similar to what one would expect to happen in the real world.

While the BCI processing and P300 detection was implemented by g.tec, our task was to develop a realistic virtual apartment including a great number of interactions as outlined above and summarised in Table 5-2 and Table 5-1.

[23] Possible selections were: TV, Music, Temperature, Movement, GoTo, Phone and Light.
[24] Common operations can be selected from several matrices and include the following 10 tasks: Main Door Camera On, Main Door Camera Off, Open Main Door, Close Main Door, ask: “Who are you?”, say: “No, thanks.”, say: “Come in.”, say: “One moment, please.”, standby P300, switch on P300.
[25] Links to previous matrix, either TV, Music, Temperature, Phone or Light.

Figure 5-5. Birds-eye view and living room of the virtual apartment.

This involved the usual steps of modelling, texturing, lighting and choosing the appropriate parameters (cf. Section 4.3.4.4), creation and definition of suitable light sources, baking the textures and exporting the meshes for use with XVR, and it was carried out in 3D Studio Max.

Since we also wanted to allow participants to switch on or off lamps, we had to bake several scene textures accounting for the various lighting conditions. For example, making one lamp interactive required two textures (i.e. one for on and one for off), while two lamps required four (i.e. both on, both off and one for each individual lamp turned on). The apartment comprised a corridor, bathroom, kitchen, bedroom and a living room and some rendered images of it are shown in Figure 5-5 above.

Regarding audio effects we played back suitable sounds in form of waveform (WAV) files whenever necessary. These mostly accompanied and underlined the visual changes. Video files representing TV channels had to be split into video (AVI format) and audio (WAV format), because XVR does not support display of bundled video content that includes audio.

Our pilot tests, especially those carried out in the CAVE, showed that many participants had problems keeping up with the rate of the P300. Since the experiment was not self-paced but task-oriented they could neither choose the order of the tasks nor the pace of the experiment and to partially compensate for this we implemented a function that allowed them to interrupt the current task. By exploiting the fact that participants wear a head tracker and knowing the rough position and orientation of the P300 display we can infer whether are looking at the display or not. This is not the case for many BCI applications and most P300-based systems struggle to offer a simple option to switch on or off the device other than through a symbol in the display itself.

In order to provide such a method we simply intersect the view plane normal (VPN) with the quadrilateral defined by the position, size and orientation of the P300 screen. If the ray and quadrilateral intersect, the person is looking at the P300 screen and otherwise away from it, possibly focussing on any part of the VE displayed on the powerwall. If the display is fixed at a certain position and angle relative to the powerwall this task is trivial, otherwise we require another 6-DOF tracker to track position and orientation of the P300-display.

This provides us with a neat method of controlling the status of the BCI: if the wearer of the head tracker is looking away from the P300 display, any current operation is aborted and the device is sent to sleep, while if he does look onto the screen, it is activated again, restarting the current task. This puts users in control of pace and activity of the device, allowing them to visually and physically explore the VE whenever they wish without being distracted by the P300. This setup is illustrated in Figure 5-6.

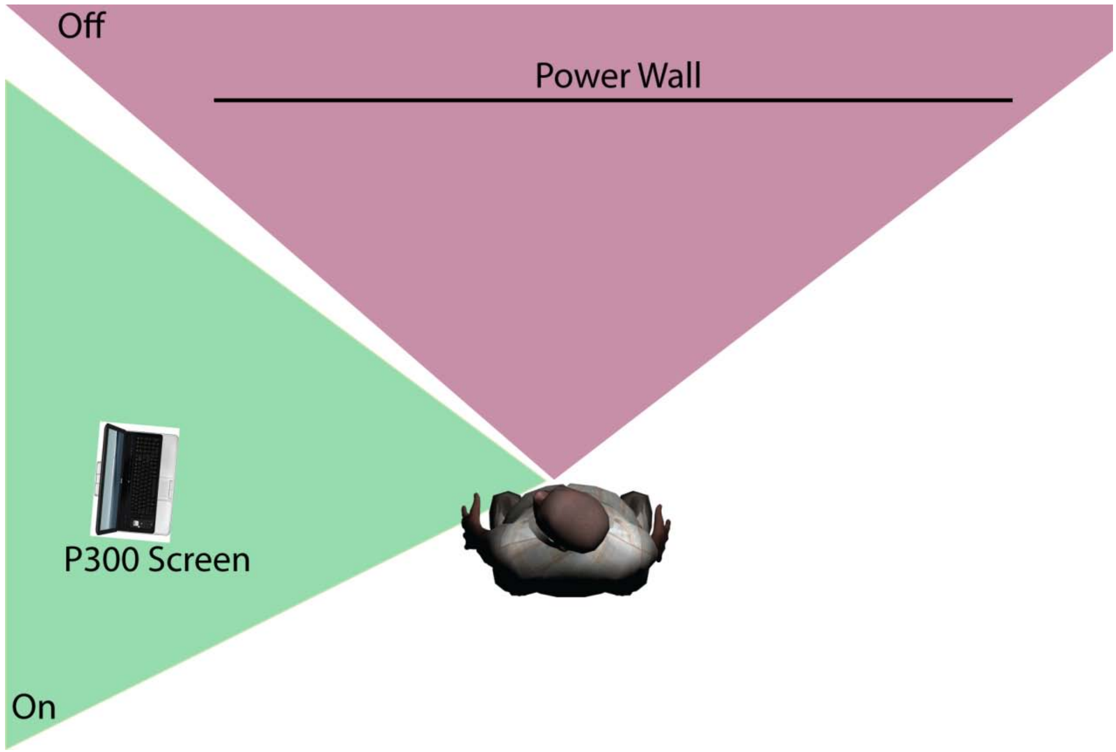

Figure 5-6. Experimental setup. When facing the P300 screen to the left starting at an angle of roughly -45º from the power wall, the P300 will activate and remain active while its user is facing in the rough direction (green area, On). When facing away from the P300 and onto the powerwall (purple area, off), the P300 interface is switched until the person points his gaze back onto the P300 screen. This ensures that he can visually and to some extent physically explore the VR without effecting undesired actions.

5.3.4 Task

Each subsequent task was displayed on the P300 prior to the next classification. The symbol the participant should focus on would appear for three seconds in a panel above the interaction matrix. There were a total of 23 tasks per condition and the order of tasks was repeated in all three conditions. The choice of tasks guaranteed that each interaction matrix was used at least once and that each room was also visited at least once. A list of symbols representing tasks is given in order in Figure 5-7.

Figure 5-7. Fixed task sequence given to participants.

5.3.5 Population

This study was conducted with 12 healthy participants, aged 20-33 (25±4 years). Two participants were female and ten were male. All of them had normal or corrected-normal vision. Although this was not a necessary prerequisite none of them had previous experience with a BCI. Participants were paid €15 for their participation. Each experiment lasted for approximately two and a half hours.

5.3.6 Procedure

Each participant was first familiarized with the equipment used: the BCI and the Powerwall plus stereo glasses. The participant was then handed an information sheet explaining the experimental setup and filled in a questionnaire on demographics that also included a section on proficiency of computer and programming skills as well as VR and gaming experience. The participant was told that he could stop the experiment at any time without giving a reason and was also asked to sign a consent form. Both forms are printed in Appendix D. The volunteer was then fitted with the BCI, which took between 15 and 25 minutes. He would then commence training. There were short breaks between each session which totalled 50 to 60 minutes per participant.

Training was necessary in order to compute a separate classifier for each of the seven interfaces and it was completed while the participant was seated in front of an ordinary desktop computer screen. This procedure was guided by the experimenters and, to begin with, the participant completed an ordinary copy-spelling task, where he or she would be presented with a word or random character sequence and was asked to spell a word sequence from the 6-by-6 matrix of alphanumeric characters. Each cycle performed 15 iterations of random highlighting of symbols before a decision was made – significantly more than in each of the later experimental conditions. A classifier was then computed and saved for the matrix presented. However it was also used as the basis for the subsequent training matrix and for each following training matrix we used the previous classifier as an input for classification which would then be adapted for the current matrix given by the participant's response to the particular training task. This way we trained a total of seven classifiers for every participant and there was one classifier per P300 matrix.

Participants were given a ten-minute break after the training session while the experimenters rearranged the setting for the experimental setup that included the powerwall. After the break the participant performed each of the three conditions in a row. Tasks were performed in the same order and at the end of a classification the symbol relating to the next task was shown on the P300 screen for a duration of three seconds before classification would start again.

For the experiment participants were asked to stand in front of the powerwall at a distance of about 1.5m. They were fitted with stereo glasses and a head tracker. They could move around freely in a space of about 2x2m in front of the wall, as illustrated in Figure 5-8 (cf. also Figure 5-6 on page 144). The P300 display was at a fixed location on the left side of the screen. In each condition, participants were asked to complete the same sequence of 23 predefined tasks. Similar to the copy-spelling task the experiment was guided rather than self-paced and each task was highlighted in advance. If the participant failed to select the correct symbol, the procedure would simply continue, so for each task the response was either correct or incorrect. The P300 interface could be activated deactivated and controlled using head movements as described in Section 5.3.3.2 above. Each condition differed only by the number of P300 decision cycles. There were eight, four and two, respectively. Note that training was carried out with 15 cycles per matrix, so the actual experiment was considerably shorter than the training period. The experiment lasted for approximately 30 minutes. The entire experiment in all three conditions including training was thus completed on the same day.

Figure 5-8. A participant concentrating on the P300 display in the space in front of the powerwall.

5.3.7 Results and Discussion

Figure 5-9 shows overall performance of the 12 subjects in each condition: 8, 4 and 2 iterations, respectively. The best result was achieved for participant 6 with 100% accuracy for 8 and 4 iterations. Worst results were recorded for participant 3 with only 30% accuracy for 2 iterations. As expected, performance was generally higher for eight iterations and became successively lower in the other two conditions. The mean performance rates for eight, four and two iterations were 79.0%, 69.6% and 53.6%, respectively. The overall performance level was 67.4%.

Figure 5-10 shows the average performance rates in each task – and matrix it appears in – and condition. All participants performed best for the move forward command (task 4, 86% accuracy) and worst for the goto command (task 20, 25% accuracy). We can also observe that while the overall trend exhibited by the mean performance rates above is reflected in each position, there appear to be two outliers, namely the two instances of the goto matrix (task 8 and 20, respectively). Results for both tasks are consistently low across all participants and all three conditions. They are much lower than one would expect considering all other tasks. One reason for this might be that the classification algorithm applied during training did not perform as expected for the goto matrix for the majority of the participants. Given that the same algorithm was applied to every matrix this is not likely to be the main cause.

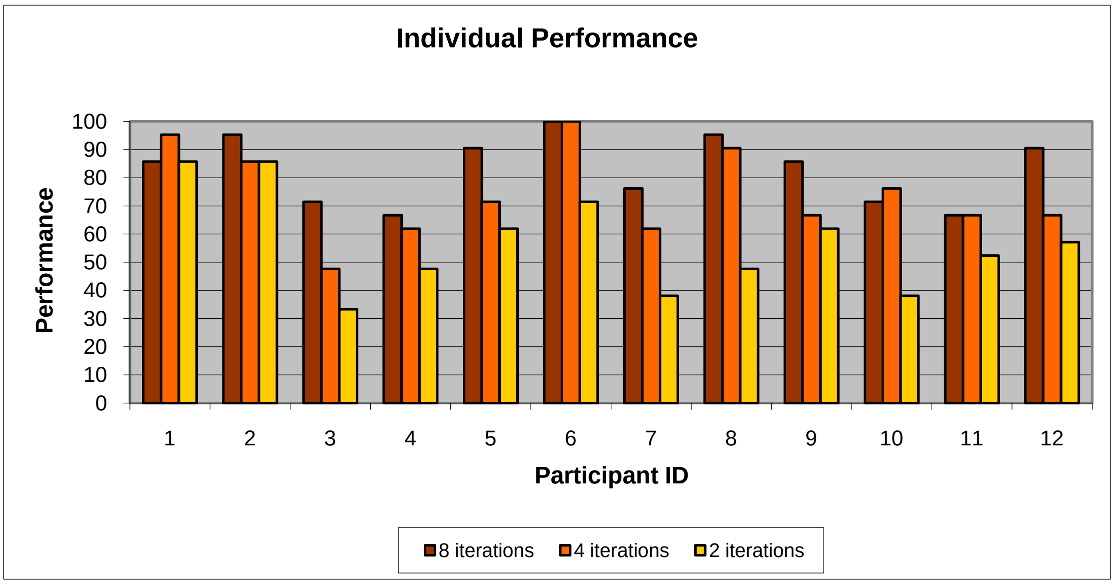

Figure 5-9. Performance of 12 participants in each condition (8, 4 and 2 iterations).

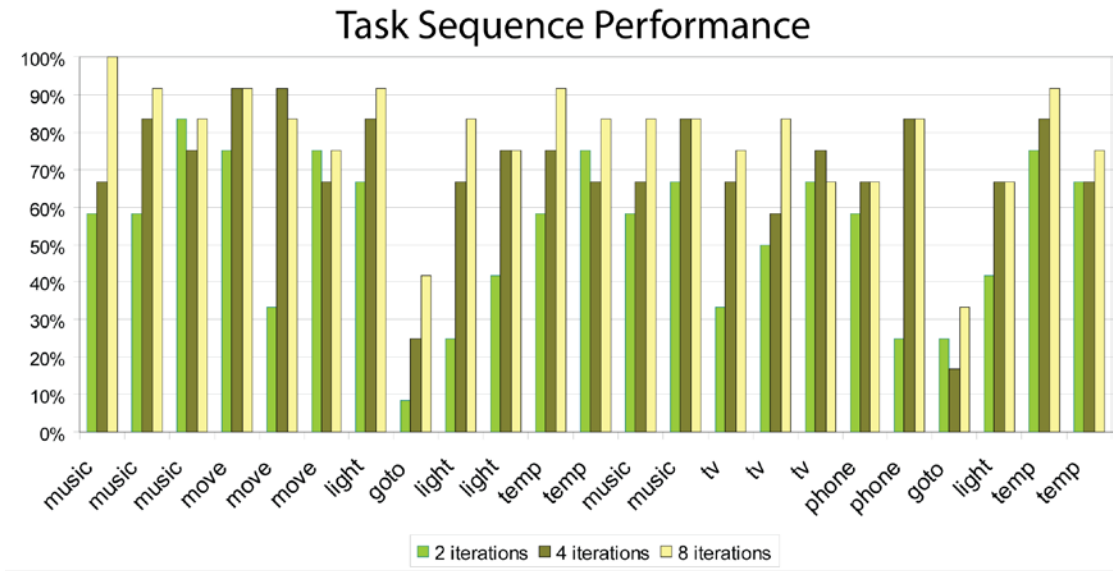

Figure 5-10. Average performance for all 23 tasks and 3 conditions.

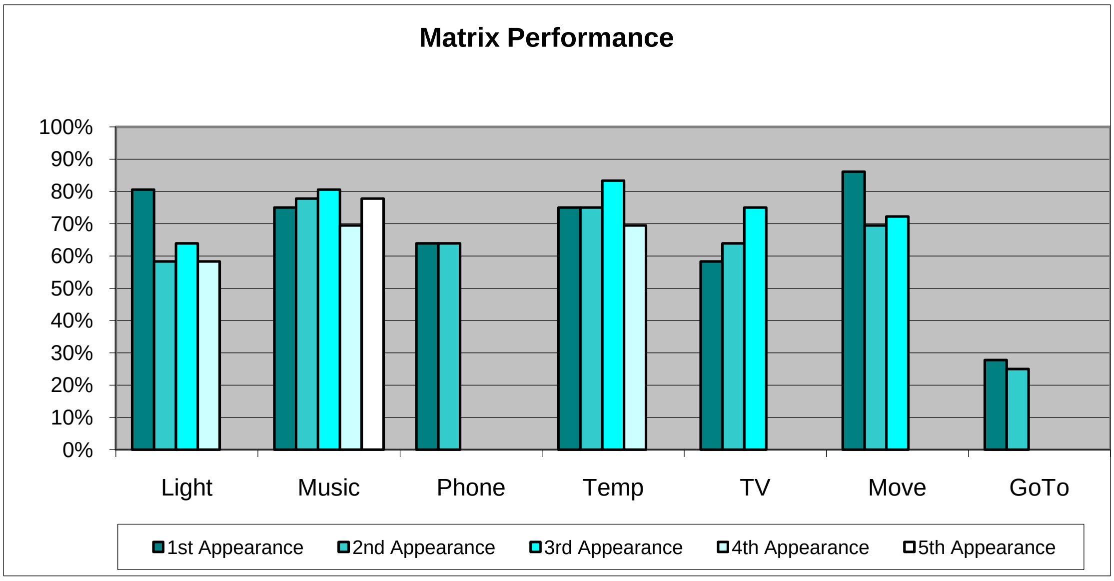

Figure 5-11. Performance for each matrix given its occurrence. Note that some matrices appeared more often than others.

Another reason could be the layout of the mask (see Appendix E) that is more cluttered and notably different from most of the others. However, the latter is also true for the movement matrix and we did not experience any significant decline in performance for tasks executed by using this matrix. It is therefore more likely that the problem arises due to the layout and there is more evidence to support this.

Figure 5-11 (above) shows accuracies for all appearances of all seven matrices. The light mask for example appeared four times and the total accuracy was 65.3%. The highest accuracies were achieved for the music, move and temp matrix, while the worst, as we already discovered above, can be found in the goto matrix.

It is worth noting that three matrices (music, temperature and movement) were controlled with around 75% accuracy. They have 50, 38 and 13 icons, respectively. Normally the amplitude of the P300 potential increases the more icons used because the likelihood is small that the specific icon is highlighted and this usually results in higher accuracy. However, the movement matrix only has 13 icons and nonetheless yields among the highest accuracy. Considering the goto mask once again, it contains 22 icons (nearly twice as many as the movement matrix) but average scores were only around 26% and the P300 analysis shows that the potential generated by the mask is smaller than for the other masks [Guger et al., 2009]. The reason for that can thus only be the layout of the mask itself and this will be investigated in further studies.

Many participants also commented on the system's behaviour. At the beginning of a new task a symbol would be randomly selected from the list of still available ones including the desired icon. Often, when this symbol appeared in the beginning, participants failed to register it, because it sometimes takes one icon flash to fully focus on the task. This means that one iteration would often be classified incorrectly. Now, for a greater number of iterations such as eight or higher, this did not matter so much, because such an incident would usually be cancelled out by the (correct) classification of other iterations. It had a much greater impact, however, the smaller the number of iterations were. If it occurred during four iterations (i.e. the second condition), a quarter of the data would be classified incorrectly without the participant getting a chance to influence its behaviour. For two iterations this effect is even more devastating because 50% of the data may be classified incorrectly as a result of this and it becomes impossible to recover from this dilemma. A simple solution for guided and possibly also for self-paced exploration exists and we will test it in future scenarios. Since all we need is the participant's attention we could flash the entire screen (or the designated symbol in guided tasks) once before the start of the first iteration. In guided exploration this would not amount to additional time taken while in self-paced exploration it could be implemented such that additional timing is minimal and that does not amount to the duration of an extra iteration.

Presence scores are presented and analysed and discussed at the end of the next sub-Section in 5.4.8 below.

5.3.8 Summary

We have presented a study evaluating usability and performance of a P300-controlled BCI as an interface to VEs. The system allows its users to control a variety of different operations through mental activity which is subsequently translated into the desired action. Our study involved interaction with a virtual smart home, which also acted as a test environment for such interaction in real smart homes. The virtual smart home is a much more flexible and cost-efficient environment than its real counterpart.

Our within-group design study showed that despite considerably short training periods participants can learn to control objects in the VE to a reasonable degree of accuracy and we were able to reach a bitrate of roughly 90bits/s.

Another achievement of this project was the design and successful operation of various P300 matrices with varying layouts. While traditional frameworks adhere to equidistant icons in a square matrix, most of our designs show that uniform distribution of icons on the matrix is not a requirement, although, as we discovered, there are some limitations to this approach.

Most participants commented that while the technology still needs improvements such a system in a real apartment could have a lot of benefits even for healthy people.

5.4 Comparative Presence Study

5.4.1 Overview and Motivations

This study differed somewhat from the previous one and our aim was to compare presence scores from the first experiment described in Section 5.3 above with presence scores taken in a separate study within the same VE. During pilots we found that many participants mainly focussed on the P300 display many without paying attention to the VR. We speculated that this might be due to the high workload associated with using a BCI and this led us to design a second study in which we could directly test this assumption. We used the same environment, though slightly modified it so that participants could interact and navigate through combination of wand and direction-of-view operations. We employed the gaze-based method because it is an action that is fairly similar to focussing on a flashing symbol in the P300 interface for some time. Instead of concentrating on symbols on a separate screen however, participants could operate virtual objects by directly looking at them.

5.4.2 Variables

Since the aim of this study was to compare the reported presence scores with those from the one presented in the previous section, the current study was confined to a single condition only. Reported presence scores are extracted from the questionnaires and used to analyze variations in the two experiences.

5.4.3 Piloting

Piloting took place over a period of one week. Since the changes to the VE were minimal we focussed on the new types of interactions. Although there was only one condition compared to the previous study, which had three, the tasks and order remained the same, so we concentrated on the participants' acceptance of the combined direction of gaze and wand interface. Gaze-based navigation and interaction techniques have been studied extensively in the past [Pierce et al., 1997, Bowman et al., 1997] so there was no need to assess its quality. The wand was only used for navigation and activation or deactivation of the gaze method.

5.4.4 Apparatus and Implementation

The basic setup and implementation was essentially the same as the one described in Section 5.3.3. However, as opposed to the previous experiment we did not use a BCI to control events in the VE, but relied on a different approach that also involved use of the wand. We explained the reasons for this in some detail in Section 5.2. The difference resulting from using a wand over the BCI was that participants were not required to switch between P300 monitor and powerwall and so could fully concentrate on the VE and thus would provide us with a good measure to probe

We thus removed the laptop and the BCI including the wireless device and replaced it with a standard handheld wand. Objects could now be activated and manipulated by looking at them. To facilitate its use we rendered a ray emanating from the participant's head position into the viewing direction. It could be activated and deactivated using a button on the wand, which was also used for navigation.

Interactive objects that intersect with the ray are highlighted by inverting their colours and activated after two seconds. An example of this is shown in Figure 5-12. Navigation is performed by using the wand so in contrast with the BCI experiment there were no teleports. The wand was also used to activate or deactivate the ray. An invisible and deactivated ray would not trigger an object. We also introduced a short training session to familiarize participants with the use of the wand and the modes of interaction. In the P300 condition we presented the symbol associated with the next task before commencing a new cycle and in the gaze-based condition the current task could be displayed as text on the screen by pressing another wand button. These were short sentences such as "*Use telephone*" that would appear on the screen while the button was pressed.

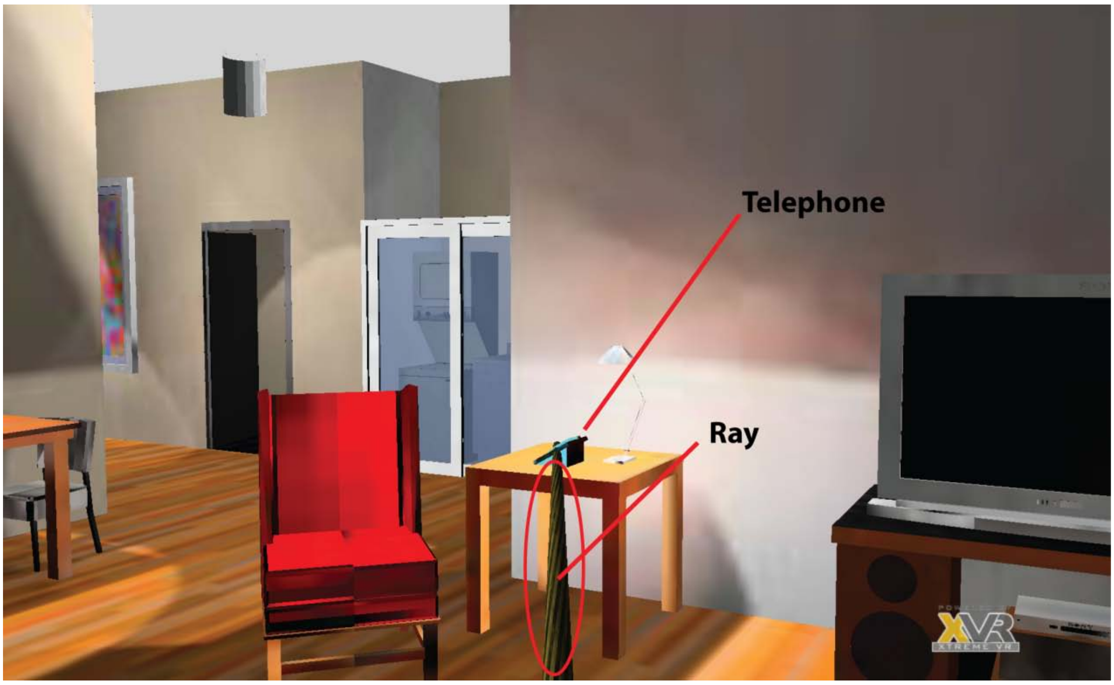

Figure 5-12. Adapted version of the smart home using gaze-based interactions. In this example the ray intersects with the telephone whose colours are inverted (original colour is red) and resting the ray on the object for a few seconds will operate it.

5.4.5 Task

Whenever possible the tasks and order given were the same as in the previous study. Teleports were translated into navigation tasks, so if a task in the BCI experiment would teleport someone into the bedroom, in this experiment the participant had to navigate there by using the wand. The same is true for simpler types of movement-related actions, such as turn right or move forward.

Recall the order of tasks given in Figure 5-7 (page 145). The first task given is “Play” – to start playing a song on the stereo system. In contrast to the BCI condition it was not possible to operate an object without seeing it; it was inevitable to slightly shift the order of some tasks. The modified task sequence is shown in the middle column of Table 5-3 on page 155. Note that implicit tasks such as finding an object and gazing at it thereby intersecting the ray with it for a few seconds until its activation are not mentioned in the table. However, it also illustrates the number of steps required for BCI-operated and gaze-based interaction. Note also that this resulted in a change in the underlying logic of the system. The BCI acts as a universal and ubiquitous device which allows its wearer to control everything from anywhere. The gaze-based method is thus more limited in the sense that only objects that are directly visible can be modified, while the carrying out tasks with the BCI requires more steps because different interaction matrices have to be selected in order to achieve a task.

Table 5-3. Overview of gaze-based task sequence and comparison with BCI operations. The number of necessary operations – except for navigation which cannot be exactly quantified – is greater for the BCI due to switching between interaction matrices and 11 out of the 23 tasks involve changing from one to another.

	Gaze-based	BCI
1	Open front door	Open front door
2	Go to living room (wand)	(a) Select 'Movement' **matrix** (b) Rapid forward (c) Turn right (d) Select 'Main' **matrix** (e) Select 'Goto' **matrix** (f) Go to location 'C'
3	Play music	(a) Select 'Music' **matrix** (b) Play
4	Toggle light	(a) Select 'Light' **matrix** (b) Toggle light
5	Switch on air-conditioning	(a) Select Temperature **matrix** (b) Switch on air-conditioning
6	Stop music	(a) Select 'Music' **matrix** (b) Stop
7	Switch on TV	(a) Select 'TV' **matrix** (b) Switch on TV
8	Switch off TV	Switch off TV
9	Use telephone	(a) Select 'Phone' **matrix** (b) Make call
10	Switch off air-conditioning	(a) Select 'Temperature' **matrix** (b) Switch off air-conditioning
11	Go to bedroom (wand)	(a) Select 'Goto' **matrix** (b) Go to location 'V'
12	Close bedroom door	Close bedroom door

5.4.6 Population

This second study was carried out with 12 healthy participants, aged 19-36 (26±5 years). Nine participants were female and three were male. All of them had normal or corrected-normal vision. Participants were paid €5 for their participation. The experiment lasted for approximately thirty minutes.

5.4.7 Procedure

The procedure was the same as described in Section 5.3.6. There was one difference relating to familiarisation with the interaction devices, which in this case were the wand and gaze as opposed to the BCI. There was an initial training environment in which participants could familiarize themselves with the wand navigation and use of the buttons. The training environment consisted of a warehouse-type building with several different-coloured cones that had to be “activated” in a certain order by intersecting the pole with the object. Participants could rely on the help of an experimenter during this training session.

5.4.8 Results and Discussion

Although our main focus was on a comparison in presence scores, for completeness the mean performance in condition 2 was 64%, almost the same as the average of condition 1, which was 67%. As it requires a lot less training than the BCI condition, the reason for this low performance may be that some tasks were somewhat ambiguous. When asked to open the bedroom door, for example, all but one participant opened the terrace door instead. This has to do with the participants not knowing the exact layout of the apartment, which was the same as in condition 1. In addition, unlike condition 1, participants had to be in line of sight of the objects and maintain certain proximity in order to trigger them. Choosing the wrong object for those reasons is therefore not a problem that arises in a BCI-type interaction, because it is not necessary to know the exact location of an object in order to trigger it. Position in virtual space and knowledge about it become largely independent of the task when using the BCI. Once an object is chosen from the list is triggered irrespective of whether the BCI user knows where it is or whether he is close by. In this sense it is a much simpler interface that requires less knowledge about the space.

Table 5-4. Means and standard deviations for presence questions in both experiments.

Question	Mean BCI	SD BCI	Mean Wand	SD Wand
Q1 To what extent did you feel like you were in the virtual apartment? (1 = not at all, 7 = most of the time)	3.0	1.64	4.5	1.88
Q2 To what extent were there moments during which you felt the apartment was real? (1 = never, 7 = most of the time)	2.92	1.51	3.92	2.15
Q3 Do you think of the apartment as an image you saw or as a place you visited? (1 = an image, 7 = a place)	2.58	1.12	4.25	2.1
Q4 During the experience did you feel you were in an apartment or in a laboratory (1 = in laboratory, 7 = in apartment)	2.91	2.0	4.83	1.85
Q5 During the experience did you think a lot you were inside a laboratory or were you absorbed by the apartment? (1 = majority of the time, 7 = hardly)	2.75	1.57	4.58	1.78

In condition 1 (i.e. BCI study) as well as condition 2 (gaze-based interaction) we asked participants to fill in a short questionnaire containing five quantitative “presence” questions. The questions (translated from Spanish) are summarized in Table 5-4 above, mean and standard deviation scores are given where applicable. Five Questions are on a 7-point Likert scale plus three questions inviting the participant to comment on specific points relating to the experience. The meaning of the extremes 1 and 7 are indicated in each question in the table above. If we take the 5 presence questions (Q1 to Q5) and compute the number of questions for which the score is greater than or equal to 5 (out of 7), we obtain a new variable y.

For condition 1: $mean(y) = 0.83$, $sd(y) = 1.53$

For condition 2: $mean(y) = 2.67$, $sd(y) = 1.83$

A non-parametric rank sum test rejects the hypothesis of equal medians (P = 0.012). If we consider each question individually then the rank sum test results in the following, shown in Table 5-5.

By examining the means and standard deviations and also taking into account these tests, it is clear that the evidence leads to the conclusion that the reported level of reported presence was higher in the second condition compared to the first.

Table 5-5. Non-parametric rank sum test for individual questions.

Question	P
Q1	0.0496
Q2	0.1663
Q3	0.0405
Q4	0.0426
Q5	0.0204

Now, regarding the subjective presence scores from the BCI study, they alone are interesting because they are overwhelmingly low. This could mean that either the workload required for operating the BCI was too high and that participants failed to register the apartment. However, about a third of the participants commented on question 6 ("How did you feel during the experience") that they liked the visual appeal of the apartment, so there is no doubt that they were aware of at least some aspects relating to its realism. One participant, though, explicitly stated that the BCI required too much visual attention. It is possible, therefore, that merely allowing participants to control the state of the BCI by looking at or away from the screen was either not a sufficiently clear procedure or switching between two different displays was too confusing. Our own observations during individual trials show however that people frequently switched back and forth between P300 and powerwall and remember also that they were located about 1.5m away from the powerwall covering almost the entire field of view when facing it directly. This implies that at least they considered the environment, though somehow the used of the P300 must have interfered with the participants' perception with respect to the VE and thus negatively affected their sense of presence.

Another possible explanation for the low scores relates to presence theory. There are some theories of presence that tend to equate action, action potential, or correlation between action and an expected and detectable outcome, with the sense of presence [Schubert et al., 1999, Flach and Holden, 1998, Zahorik and Jenison, 1998]. In light of the current study, this may be the case if and only if the action is effected by means of at least some physical activity. Be this activity based on mere button presses or, at the other end of the spectrum, more physically engaging approaches, is irrelevant because compared to interaction using a BCI most of these depend on a person's physical activity while the BCI is a purely mental procedure. Thus, one reason for the low scores may be the unusual and unfamiliar method of communication compared with more physical means. Some comments point in this direction and one participant stated that "*It's weird to realize something […] without any physical interaction. I felt like I was missing something*". However, in a previous report where the objective was to move a virtual body by thought by using motor imagery participants reported the opposite and that the experience became more dreamlike [Friedman et al., 2007].

A fairly novel mode of interaction that uses only thought, therefore, may appear too vague in many aspects and perhaps bizarre. To some extent there is neuroscientific evidence supporting this view and some work demonstrates that a substantial part of our self-perception and recognition is obtained from action [Rochat, 1998, van den Bos and Jeannerod, 2002], which is physical in nature and possible there is simply not enough correlation between the physical action and the process of executing it, i.e. the action is not imagined but achieved by counting repeated occurrences of a symbol representing that action. Interpreting the results in terms of sensorimotor contingencies we can say that from the point of view of the BCI user there are no VEAs because unlike physical actions that may have previously been learned, such as moving the mouse to the left in order to move the cursor on the computer screen to the left, using a BCI completely lacks physicality and may thus not be registered as a sensorimotor skill because it does not involve motor activity. On the other hand, a recent study on inducing the rubber hand illusion through motor imagery showed that body ownership was produced in many participants [Perez-Marcos et al., 2009] with similar results to the original study [Botvinick and Cohen, 1998]. However, motor imagery is a much more active type of BCI than the rather passive and responsive P300 interface and thus may be more similar to actual physical action than the use of the P300.

5.4.9 Summary

In this section we presented a complementary study assessing the effects BCI use has on presence. We measured presence in two task-oriented studies with different interaction methodologies but otherwise comparable setup. One uses a BCI for interaction and another one a gaze-based selection approach, which we deemed sufficiently similar to the P300-based interface of the BCI to allow us to compare presence scores between both conditions.

Quantitative presence scores show that the self-reported sense of presence was significantly higher in the second condition than in the first one. Principal reasons for this may be workload and attention but also the experimental setup that did not permit the participant to view the P300 interface and VE at the same time.

5.5 Chapter Summary

In this chapter we presented two experiments. In the first one, we evaluated the performance of a P300-based BCI system connected to a VR system. The VE consisted of a virtual smart home environment with over 200 commands. We conducted a within-groups study with 12 participants in three conditions and our results suggest that the BCI is capable of delivering high performance and accuracy (~67%). We also showed that this still holds despite very short training times – about one hour instead of several days or weeks – so the BCI is a well-suited interface for control of smart homes and VEs in general. Results were generally high given the number of classification iterations, 8, 4 or 2.

For this study we also implemented and tested a new set of P300 matrices that do not adhere to previous layout rules where any symbol and not only alphanumeric or language-specific characters can be aligned in almost any fashion as opposed to a rectangular grid. This is an important result because it shows that arbitrary symbols can be used for this task and that they can be arranged, though possibly with some constraints, in virtually any fashion (cf. Section 5.3.7).

Mental workload, the experimental setup or both might be responsible for overall low presence scores in condition 1. A mean score of 2.8 over all items in the presence questionnaire, which is considerably lower than typical results, suggests that one or both may have constrained participants from perceiving the VE as real. In condition 2 the overall mean presence score was 4.4.

In the second study in a single condition we directly compared the BCI setup with a more traditional interaction method based on a combination of gaze and wand-based operations. The environment and setup remained the same as in the initial experiment. While overall performance rates are very similar in both conditions (i.e. approximately 65%), the reported level of presence was significantly higher in the second study than in the first one. We conclude that this is the result of mainly two issues. First, rather than the experimental setup itself (cf. Figure 5-6) high mental workload in the BCI condition possibly inhibits people to *willingly suspend disbelief*.

Second, the lack of physical action especially during navigational tasks was so unusual and novel to all participants that it resulted in a decisive lack of physicality when viewing the VE. Since all participants were healthy they quickly found that their physical movements had little to no effect, which might be an unfamiliar experience to most and possible causes were discussed in Section 5.4.8. The general problem could relate somehow to the fact that BCI navigation completely lacks physical activity. Tasks in the BCI condition could be completed independent of any knowledge about the virtual space, which was vital for completing the gaze-based condition.

In the second condition, participants used the wand for navigation. While there is not much physical activity triggering motion either there is a strong correlation between hand and finger activity and virtual movement, something which is completely absent in the first condition. In addition, head movements in the first condition also had the function of enabling or disabling the P300 interface – in the second condition they only gave rise to perceptual changes in the VE. Also, current software and hardware limitations prevented us from accurately implementing the P300 flashing interface directly within our VR scripting environment XVR, which may also partially contribute to this problem. Since the classification algorithm is very sensitive to small temporal disparities we were unable to integrate it into the VE and overcoming this problem might effectively and noticeably increase reported level of presence.

Regarding future work, instead of using an external handheld device or computer screen to focus on it would be useful to fully integrate the P300 paradigm within the VE. This would allow for a new dimension of P300-based BCI to be tested since the control would be completely enclosed into the VE.

6 Conclusions

6.1 Overview and Key Questions Addressed

One of the chief attractions of VR lies in the ability for 3D spatial interaction in a highly immersive environment. The biggest research challenges in this regard are the integration of multisensory stimuli and support of natural actions when interacting with a VE. These two phenomena have been discussed extensively in the presence literature and, more recently, they have been identified as some of the key aspects of making presence happen. An investigation revolving around any of these two topics is no doubt valuable to further or confirm existing theories of presence. In this thesis we developed three projects to explore and evaluate three distinct types of whole-body interfaces and test them in terms of plausibility and place illusion, two elements of a new approach to presence that we presented in Section 2.4 [Slater, 2009]. Place illusion (PI) is the illusion of being physically located in a synthetic environment and, orthogonally, plausibility (Psi) is the illusion that events happening in the VE are really happening. Two projects, i.e. the ones presented in Chapters 3 and 4 dealt with Psi and how behaviour and actions can be used to influence processes in a VE. In the last two studies presented in Chapter 5 we attempted to shed light on how radically different means of interaction used within the same environment can affect PI. Each project refers back to one or more of the central research questions concerning plausibility and place illusion we posed in Chapter 1 (cf. page 25).

> *1. Can intentional physical actions and coherent responses in the VE lead to an enhancement of plausibility?*
>
> *2. Can unintentional physiological processes enhance the experience and affect plausibility?*
>
> *3 and 4. Are BCIs useful as an input device for VEs and do how do they affect place illusion?*

6.2 Review of Experiments

The SPIN_off project introduced in Chapter 3 dealt with whole-body movements that were translated into virtual energy. This energy was then used to change the actions of a collection of virtual agents that displayed low-level intelligence and behaviour at a collective and individual level. Our hypothesis was that more action resulting in an increased response of the VE would lead to a greater connection between human and environment. Our findings indicate that one driving factor for experiencing plausibility is consistency between input action and expected outcome which supports the theory that the idea of SCs can be extended to be used for VE as well. If there is no consistency between actions and sensory responses then Psi does not arise or is much lower than otherwise. While in reality most humans experience Psi most of the time, the absence of Psi may occur only in people with neurological illnesses, but in VR it may happen more frequently.

In the Physiological Mirror, the second study on plausibility covered in Chapter 4, we introduce a slightly more subtle and unusual approach to interaction in the sense that the set of input parameters cannot be readily accessed and altered by a person, although they are produced by them. They even combine some unconscious internal reflections on exterior events. We used three types of physiological measurements – i.e. heart rate, respiration and galvanic skin response – to visually reflect the participant's physiological, and to that end psychological, state inside the VE. In our experiments we were interested in whether humans were able to detect the "real" part, a function of their own bodily state, in the VE and thus be able to form a deeper connection with either that part or the entire environment. Clearly, what happened was that there were no obvious VEAs as they consisted of permanent and unconscious bodily events that most people can not readily access or even control. Merely providing sensory variations in accordance to some "action" effected by the human may therefore not be sufficient to produce Psi. Finally, even though we were not able to confirm our hypothesis that these unintentional physiological events visualized inside the VE would strengthen the bond between a participant and a virtual character reflecting on those physiological states, we gained a lot from modifying the experimental setup intending to find the right metric for a display metaphor of physiological processes and possible obstacles. We will revisit these issues in more detail in Section 6.4 below.

In two more studies covered in Chapter 5 we described two experiments where we evaluated brain-computer interfaces for use in VEs and evaluated performance and self-reported presence scores with a complementary form of interaction based on gaze. In our particular example, the BCI was used to operate over two hundred household items and objects and navigate through a virtual apartment while gaze was used in the second study to achieve the same plus wand navigation. While the focus of the first of these studies was to assess the feasibility and performance of BCI use for VR, the overarching goal of the two studies combined was to evaluate how a P300-based BCI affects place illusion. This second part of the study was motivated by our own observations throughout the first study as well as reports on high mental workload when using such a BCI. Regarding the latter two studies involving the P300-based BCI, the results clearly showed that the self-reported presence was much lower when using a BCI. We concluded that the device must somehow inhibit its wearer's perception of exterior events and possibly introduces breaks in presence.

6.3 Main Contributions

Research on whole-body interaction is arguably still at an early stage, especially as a tool for presence research. This thesis made the following contributions. In a critical literature review we gave a detailed account of the history of presence research indicating that the terminology used has become vague and justified why there is a need for a fresh start. We presented a more recent concept of presence and distinguished it with older theories. Furthermore, all the work presented in this thesis relates back to terminology defined within this model.

We also compared and equated terminology from Slater's recent concept on presence with others used in the user interface design and the whole-body interaction community. This is an important accomplishment since being able to relate concepts in presence to those from related fields, assumptions and assertions made throughout this thesis gain a more universal character. Also, our claim is that our work is relevant not only to presence researchers but also of interest to other communities and the ability to equate terminologies allows us to place it in different contexts.

Our main empirical contributions directly reflect on the key questions we initially posed. The first question essentially asked whether it valuable to apply the theory of sensorimotor contingencies to presence research and if a sensorimotor skill coupled with co-occurring sensory variations would produce Psi. Our findings from the SPIN_off project support this hypothesis and that it still holds even if the sensory variations were previously unknown.

The second question was similar to the first and also addressed Psi, but instead of focussing on a sensorimotor skill, physiological processes, that we are often unaware of, were used. In the Physiological Mirror we were not able to demonstrate this but an important finding of this research may simply be that only conscious sensorimotor activities can produce Psi, although this may contradict with similar findings from neuroscience [Rizzolatti and Craighero, 2004] and studies on mimicry in social psychology [Grahe and Bernieri, 2002, Lakin and Chartrand, 2003].

Questions three and four related to PI but in some way were also similar to question two. They concerned the relationship between BCI in a VE and the incidence of PI and we were able to show that – somewhat similar to the second study on unconscious physiological activity – excess of mental activity inhibits PI. We demonstrated that several hundred items can be controlled in a VE by using the BCI. Our work is also valuable for the BCI community because we established that non-uniform P300 matrices can be used instead of evenly-distributed symbols in a square matrix (cf. Appendix D). Also we were able to dramatically reduce training times for several interaction matrices to a few hours at most.

6.4 Directions for Future Work

Regarding the general framework of whole-body interaction within presence in terms of PI, Psi and the body, a lot of work remains to be done. There is a clear need to integrate multiple sensory modalities such that the body – the point of fusion of PI and Psi – offers more sensorimotor skills and thus facilitates agency and coherent SCs. This is irrespective of whether one uses his real body in VR or a virtual one with some degree of ownership.

Regarding presence we have shown how different devices and interaction metaphors can influence either PI or Psi. It would be interesting to combine these findings and see how they relate to both at the same time. For example, does the use of a BCI result in low PI but high Psi or do arbitrary types of interaction based on body movements result in high PI and Psi at the same time? What does this tell us about how we perceive our own bodies versus our brains or minds and how can we use this information to improve VR?

One of the more interesting and recent crossovers of research lines no doubt is the intersection of BCIs with VR. In Section 2.6 we saw many interesting examples of BCI use that go beyond rehabilitation which in itself is a valuable field of study; and we also presented our own system in Chapter 5. The main reason why BCIs are not used beyond research and even in research only within a small domain is because their frame rate is too low and their use is tedious, tiring and frustrating compared with more traditional approaches. And these are attributes which are certainly true for types of BCI other than the P300-based approach. So why use them in VR? VR has a solid history in rehabilitation that goes beyond its recent coupling with BCIs so combining the two. Regarding less restricted application domains in VR that use a BCI, however, BCIs first need to be able to offer adequate processing frame rates that are comparable to other means of interaction. For healthy people, a joystick, wand, keyboard, mouse or combination of them is infinitely more convenient because desired tasks can be achieved in an instant. They may be crude like using the wand for navigation, for example, but they work well. Other problems that need to be addressed are training, classifier robustness and also the method of choice (e.g. P300, VEPs, SCPs, motor imagery). Training times need to be dramatically reduced to a tolerable level with some BCIs taking weeks or months to train one person how to use it properly. This is in particular why the P300 interface is currently very popular because it is simple to use and train and it yields sufficiently good results in most people. Similar to VEPs, however, they work by presenting a known sensory, typically visual, stimulus to the BCI user who merely responds to it. This may appear to be interactive but in essence it is an extremely restricted way of dictating a very limited number of possibilities because none of the actions presented by the stimuli are really scalable especially when it comes to finer and more gradual activities such as locomotion. This is almost akin to comparing early text-based computer games that had only a limited vocabulary that had to be acquired during game play, with modern games offering players a plethora of possibilities including playing with other players online. If BCI research can find more powerful and faster methods than these or radically improve them and if general health and safety concerns regarding for instance ECG implants are met, they may one day become the ultimate tool for interaction. It is almost certain that BCIs will take this leap in the near future, but currently they are too limited to compete with other interaction devices available for VR.

Our findings from the SPIN_off project essentially tell us that body movements and correlated events in the VE can enhance plausibility. As this project dealt with rather unusual movements and an abstract environment it would be interesting to create a more realistic scenario in which natural hand gestures such as grasping and manipulating an object, for example, are used. Indeed, more realistic settings may not only enhance plausibility but also be applicable to domains other than entertainment.

In our view the most challenging and exciting experiment presented in this thesis was the Physiological Mirror. Since, to our knowledge, no attempt has been made to use physiological measurements in this way of emphasizing a virtual character, it remains a difficult but interesting problem to solve. We do believe that it is generally possible to enhance the connection or general feeling between a real person and a virtual character or event by means of real physiological data coming from that person. However, it seems to be a matter of finding the right metric of physiological measurements displayed at the right intensity within a framework of suitable and easily understandable metaphors. This entails that we must first identify which physiological measurements can be easily translated into a metaphor. The use of GSR, for example, appears to be somewhat erratic because there is no tangible reflectional state that can be associated with it other than events one responds to such as arousal, excitement and so on. Respiration clearly is controllable and perceptible and so is heart rate for some people. More conscious physiological activities such as EMG might also be appealing for further study. Finding a good metaphor for each of the physiological states is the second aspect that is of interest here. Clearly, we are interested in finding the boundaries and possible circumstances under which a notable change of perception takes place for the majority of people under roughly the same conditions. If, for example, physiological feedback can one day be used in order to intensify and emphasize an interactive narrative, then it could be used as a compelling tool to drive a story, possibly inducing and generating emotional states with respect to certain scenes or characters. Something that is akin to a musical score in modern movies yet more tailored to one’s own reactions.

This thesis has aimed at developing and evaluating whole-body interfaces for the enhancement of presence in terms of place illusion, plausibility and the body. Our findings suggest that direct physical interaction positively affects plausibility while purely mental interaction via a BCI has a detrimental effect on place illusion. Our results from experimenting with physiological interaction, on the other hand, suggest that it is too difficult for people to even discern their own visualized physiology from others and that it thus cannot be used as imagined in order to enhance plausibility. Future work will no doubt build on these findings possibly by balancing the use of physical and mental interaction possibly including physiological feedback and combined they may yield a richer and more realistic experience of virtual environments.

Appendix A: Questionnaires used in the Spin_off experiments

1. How responsive was the environment, presented to you on the screen, to actions that you initiated?

Not at all 1	2	3	4	5	6	Very responsive 7

2. How much were you able to control the environment?

Not at all 1	2	3	4	5	6	To a great extent 7

3. How natural did your interactions with the environment seem?

Very unnatural 1	2	3	Neither natural nor unnatural 4	5	6	Very natural 7

4. It seemed as if my actions did not have any impact on the development of the environment.

Strongly Disagree 1	2	3	Neither Agree nor Disagree 4	5	6	Strongly Agree 7

5. How much did the visual aspect of the environment involve you?

Not at all 1	2	3	4	5	6	To a great extent 7

6. How much did the auditory aspect of the environment involve you?

Not at all 1	2	3	4	5	6	To a great extent 7

7. How much did your experiences in the environment seem consistent with your real-world experiences?

Not at all 1	2	3	4	5	6	A great deal 7

8. Were you able to anticipate what would happen next in the response to the actions that you performed?

Not at all 1	2	3	4	5	6	A great deal 7

9. How involved were you in the experience?

Not at all 1	2	3	4	5	6	A great deal 7

10. How strong was your sense of interaction with the environment?

Very Weak 1	2	3	Neither Weak nor strong 4	5	6	Very Strong 7

11. To what extent if at all did you feel you were in control of the objects and the environment?

Not at all 1	2	3	4	5	6	A great deal 7

12. To what extent if at all did other people have an impact on the objects?

Not at all 1	2	3	4	5	6	A great deal 7

13. Did you establish some form of group effort in order to interact with the objects?

Strongly Disagree 1	2	3	Neither Agree nor Disagree 4	5	6	Strongly Agree 7

14. How much time did you spend in the space?

___________ minutes.

15. How aware were you of events occurring in the real world around you?

Not at all 1	2	3	4	5	6	A great deal 7

16. To what extent if at all did you relate to or identify with the objects?

Not at all 1	2	3	4	5	6	A great deal 7

17. To what extent did you feel the objects were linked to you?

Not at all 1	2	3	4	5	6	A great deal 7

18. To what extent if at all did you feel that there were times at which the environment became the dominant reality for you and you almost forgot about the surroundings of the real space?

Not at all 1	2	3	4	5	6	A great deal 7

19. To what extent if at all did you feel that there were times at which the artificial environment and real space merged into a hybrid space?

Not at all 1	2	3	4	5	6	Very often 7

20. To what extent if at all did you experience that you were in the same place as the objects?

Not at all 1	2	3	4	5	6	A lot 7

21. Please write any other statements about what you felt or perceived during this experience.

Appendix B: Questionnaires used in the Physiological Mirror

1. On a scale from 1 to 7, where 7 represents your normal experience of being in a real place, please write down your sensation of being in the virtual room[26].

I had the sensation of being in the virtual room….

Not at all 1 2 3 4 5 6 7 almost all of the time

2. To what extreme were there moments during the experience in which you perceived the room as real?

There were times during which the experience in the room was real....

never 1 2 3 4 5 6 7 almost all of the time

3. When you think about the experience do you think about the virtual room as an image that you saw or a place that you visited?

The room appeared like.....

An image that I saw 1 2 3 4 5 6 7 a place that I visited

4. During the experience what sensation was stronger, your sense of being in the laboratory or of being in the virtual room?

I had the strong sensation of being…..

in the laboratory 1 2 3 4 5 6 7 in the room

5. During the duration of the experience did you often think of being in the laboratory or were you absorbed by the virtual room?

During the experience I thought that I really was in the laboratory….

The majority of the time 1 2 3 4 5 6 7 hardly ever

6. What do you think of the experience? How did you feel during the experience?

[26] All questions translated from Spanish.

Appendix C: Questionnaires used in the Smart Home experiments

1. On a scale from 1 to 7, where 7 represents your normal experience of being in a real place, please write down your sensation of being in the virtual apartment.[27]

I had the sensation of being in the virtual apartment….

Not at all 1 2 3 4 5 6 7 almost all of the time

2. To what extreme were there moments during the experience in which you perceived the apartment as real?

There were times during which the experience in the apartment was real....

never 1 2 3 4 5 6 7 almost all of the time

3. When you think about the experience do you think about the virtual apartment as an image that you saw or a place that you visited?

The apartment appeared like.....

An image that I saw 1 2 3 4 5 6 7 a place that I visited

4. During the experience what sensation was stronger, your sense of being in the laboratory or of being in the virtual room?

I had the strong sensation of being…..

in the laboratory 1 2 3 4 5 6 7 in the room

5. During the duration of the experience did you often think of being in the laboratory or were you absorbed by the virtual apartment?

During the experience I thought that I really was in the laboratory….

The majority of the time 1 2 3 4 5 6 7 hardly ever

6. Did you notice a difference between the three tests?[28]

7. What do you think of the experience? How did you feel during the experience?

8. If the BCI was the only way of communication in your house how useful and usable would you find it?

[27] All questions translated from Spanish.
[28] Questions 6 and 8 omitted in the second condition.

Appendix D: Demographic and Consent Forms used for Physiological Mirror and Smart Home Experiments

DEMOGRAPHICS FORM[29]

<table>
<tr><td>Your given ID number</td><td></td></tr>
<tr><td>Your Age</td><td></td></tr>
<tr><td>Your Gender</td><td></td></tr>
<tr><td>Occupational Status</td><td>Undergraduate Student
Masters Student
PhD Student
Research Assistant/Fellow
Staff – systems/technical
Faculty
Administrative Staff
Other, please indicate:</td></tr>
<tr><td>Are you taking any medication?</td><td>YES/NO. If YES, please specify:</td></tr>
<tr><td>Did you consume more than 2 units of alcohol within the last 6 hours?</td><td>YES/NO</td></tr>
<tr><td colspan="2">Please state your level of computer literacy:

(novice) 1 2 3 4 5 6 7 (expert)</td></tr>
<tr><td colspan="2">Please rate your level of experience with computer programming:

(novice) 1 2 3 4 5 6 7 (expert)</td></tr>
<tr><td colspan="2">Have you ever experienced “virtual reality” before?

(no experience) 1 2 3 4 5 6 7 (extensive experience)</td></tr>
<tr><td>How many times did you play video games (at home, work, school or arcades) in the last year?</td><td>Never
1-5
6-10
11-15
16-20
21-25
>25</td></tr>
<tr><td>How many hours per week do you spend playing video games?</td><td>0
<1
1-3
3-5
5-7
7-9
>9</td></tr>
</table>

[29] Document translated from Spanish

INFORMED CONSENT FORM[30]

Project: PRESENCCIA
Investigators: Mel Slater, Christoph Groenegress

To be completed by volunteers.
We would like you to read the following questions carefully.

Have you read the information sheet about this study?	YES/NO
Have you had an opportunity to ask questions and discuss this study?	YES/NO
Have you received satisfactory answers to all your questions?	YES/NO
Have you received enough information about this study?	YES/NO
Which investigator have you spoken to about this study?	

Do you understand that you are free to withdraw from this study?

At any time	YES/NO
Without giving a reason for withdrawing	YES/NO

Do you understand and accept the risks associated with the use of virtual reality equipment?	YES/NO
Do you agree to take part in this study?	YES/NO
Do you agree to be video taped?	YES/NO
Do you agree to be audio taped?	YES/NO
Do you agree to be physiological monitored?	YES/NO

I certify that I do not have epilepsy.

I certify that I will not be driving a car, motorcycle, bicycle, or use other types of complex machinery that could be a danger to myself or others, within 3 hours after the termination of the study.

Signed…………………………………...…………**Date**……………………………...

Name in block letters.…………………………..…………………………………...

Investigator………………………………………………..………………………..

In case you have any enquiries regarding this study in the future, please contact:

Mel Slater

Facultat de Psicologia. Universitat de Barcelona Campus de Mundet - Edifici Teatre

Passeig de la Vall d'Hebron, 171
08035-Barcelona

Tel +34 93 403 9618
Fax +34 93 402 1362

m.slater@gmail.com

Information that we collect will never be reported in a way that individuals can be identified. Information will be reported in aggregate, and any verbal comments that you make, if written about in subsequent papers, will be presented anonymously.

[30] Document translated from Spanish.

Appendix E: Seven P300 Matrices used in Smart Home Experiment

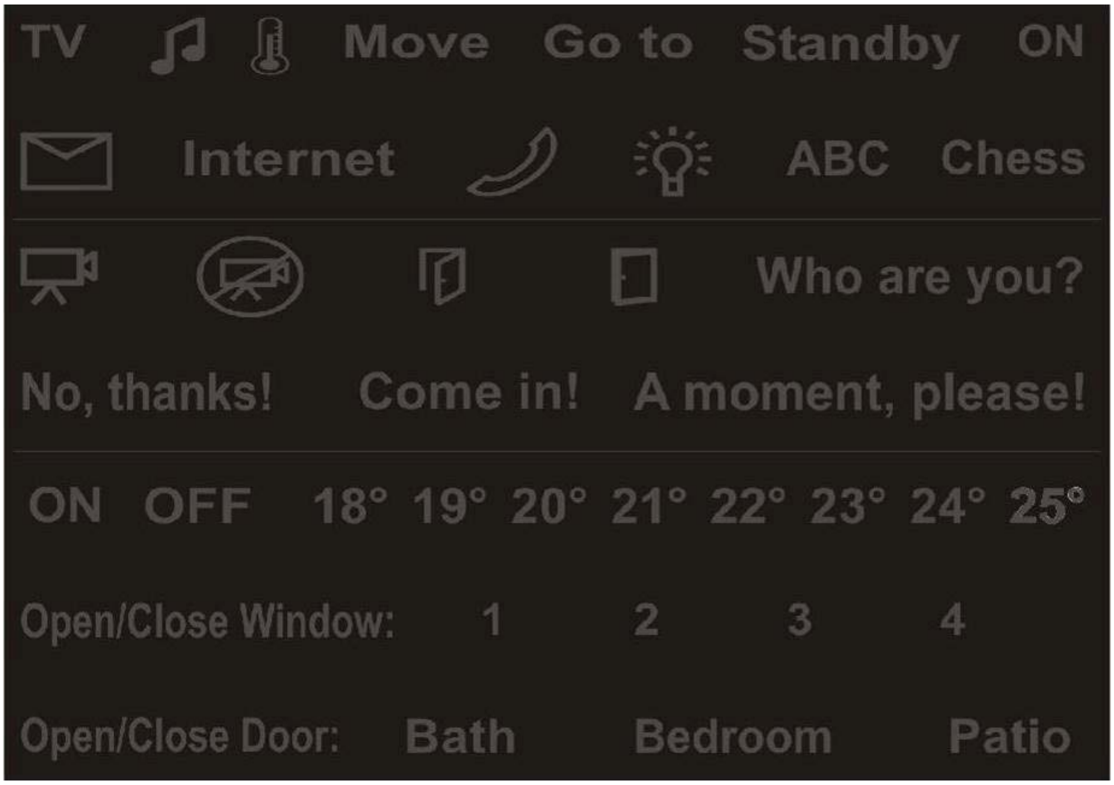

Figure E-1. Temperature matrix with 30 possible operations.

Figure E-2. GoTo matrix with for teleporting. It contains 22 symbols.

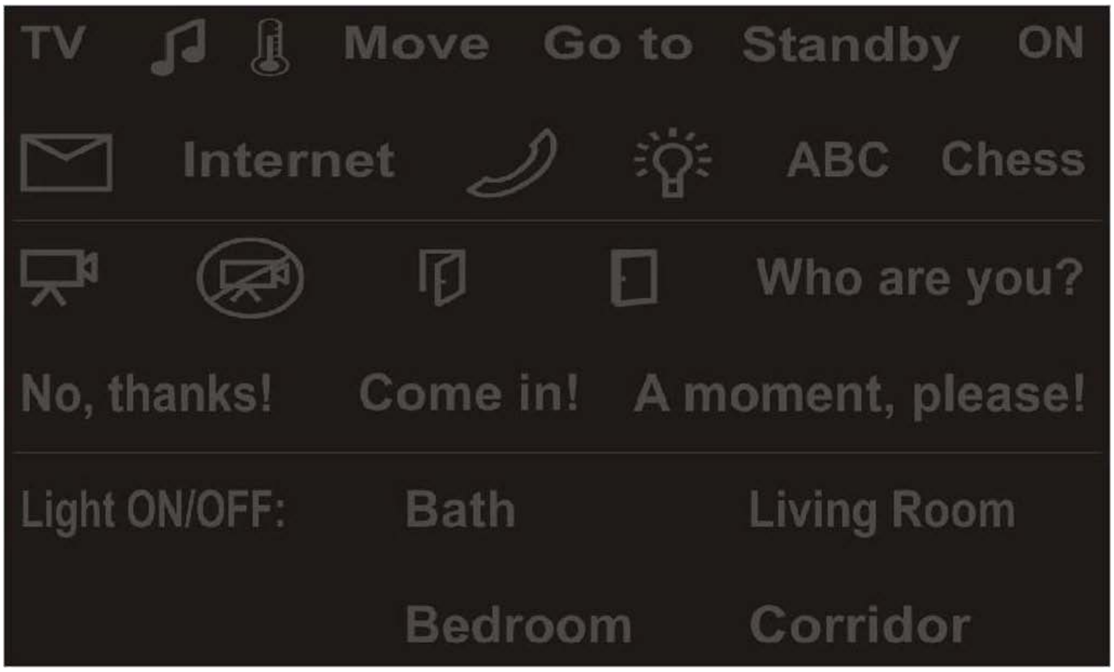

Figure E-3. Light matrix with 18 possible operations.

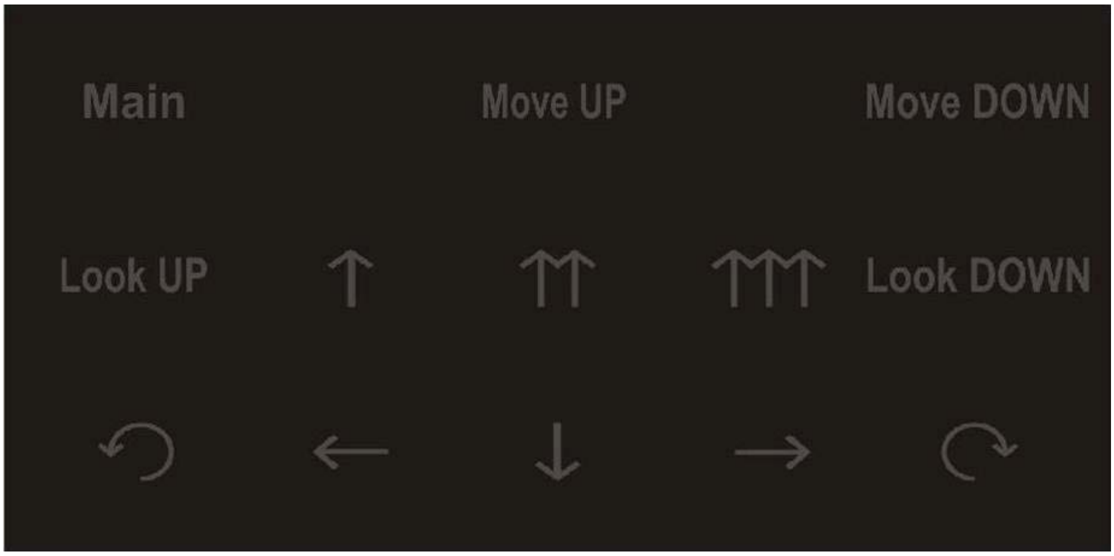

Figure E-4. Movement matrix with 13 operations.

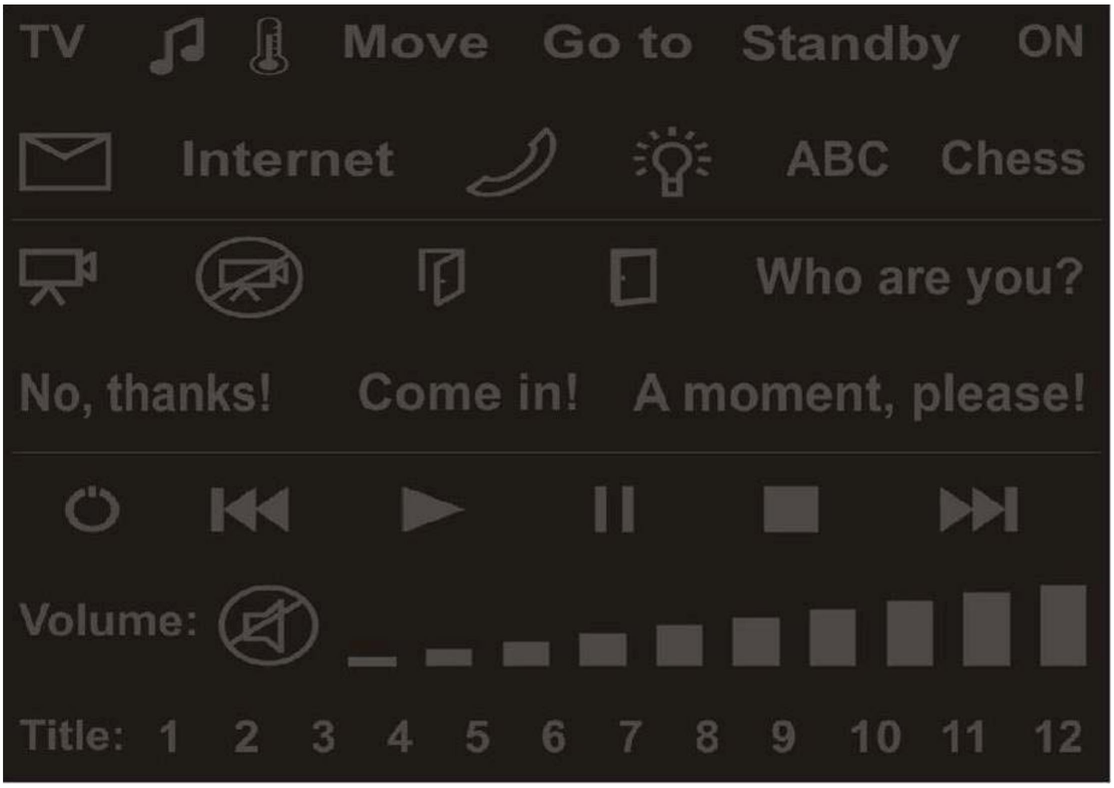

Figure E-5. Music matrix with 46 operations.

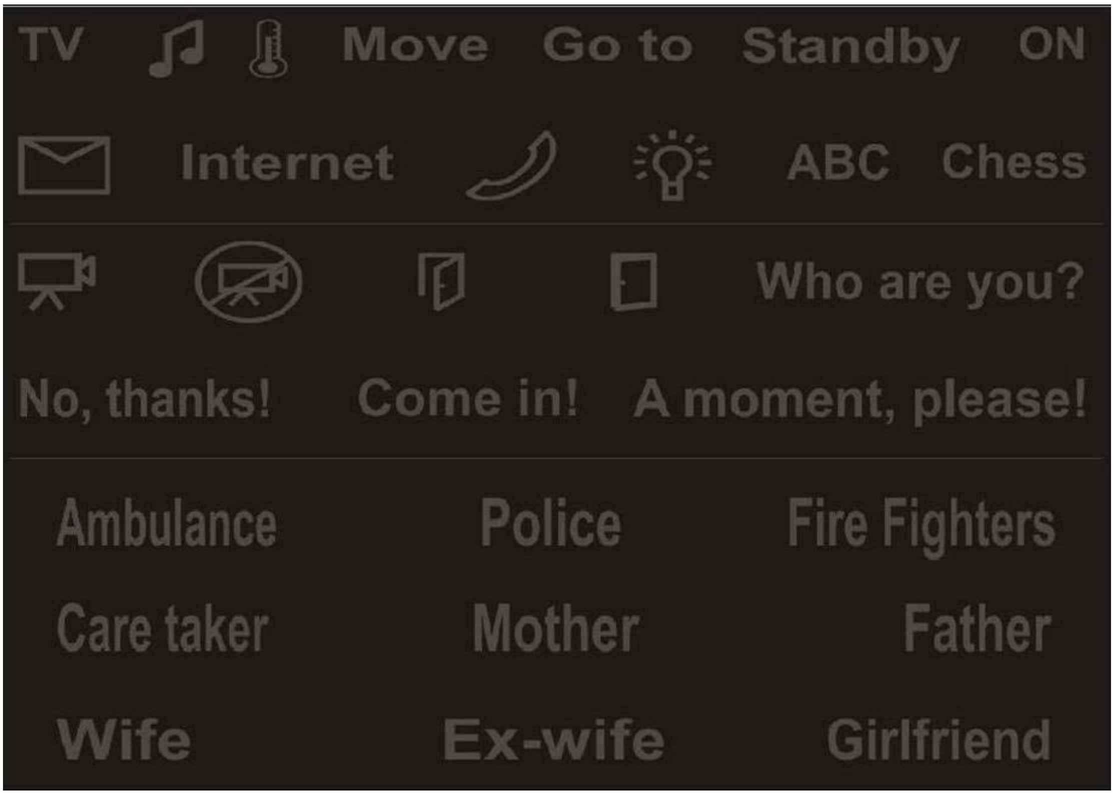

Figure E-6. Phone matrix with 26 possible operations.

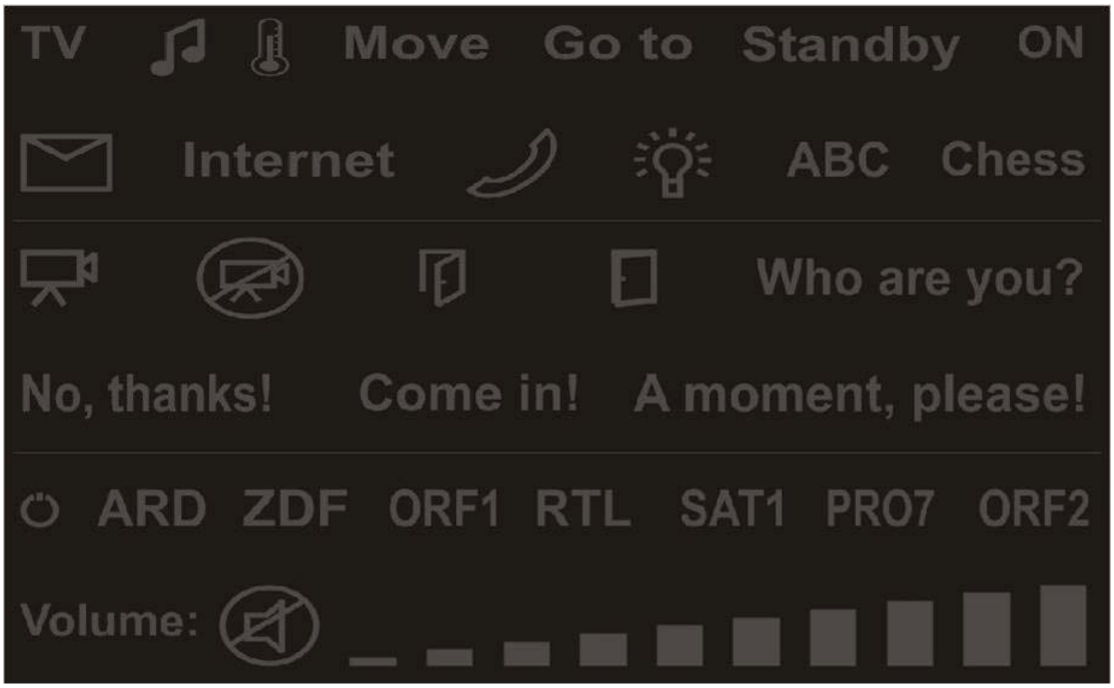

Figure E-7. TV matrix with 36 operations.

Appendix F: Simulink Models for Physiological Processing

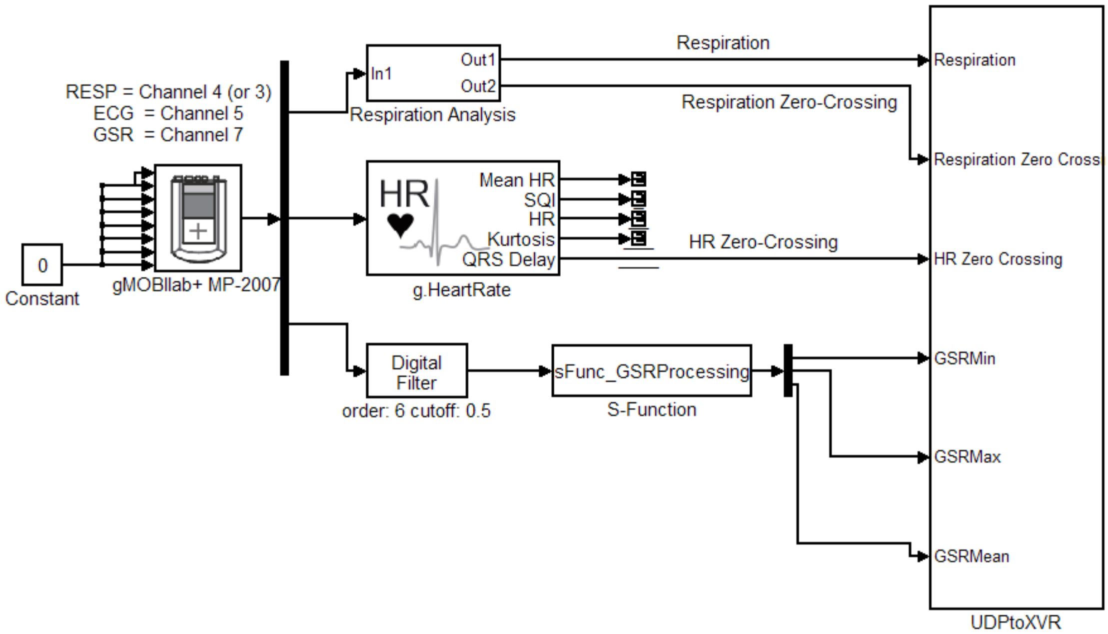

Figure F-1. Simulink model developed for the Physiological Mirror. The leftmost block titled gMOBIlab+ MP-2007 is the source where the raw signal is split up into respiration heart rate and GSR. For each of these a filter and analysis process exists in the middle. The data are combined again after analysis and transferred to the XVR program via UDP (right).

References

[Ahmed et al., 2005] Ahmed, N., de Aguiar, E., Theobalt, C., Magnor, M., and Seidel, H.-P. (2005). Automatic generation of personalized human avatars from multi-view video. In *VRST ’05: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 257–260, Monterey, USA. Association for Computing Machinery (ACM), ACM.

[Alais and Burr, 2004] Alais, D. and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14:257–262.

[Andreassi, 2000] Andreassi, J. (2000). *Psychophysiology human behavior and physiological response*. Publisher: Erlbaum Associates.

[Antley and Slater, 2009] Antley, A. and Slater, M. (2009). The effect on lower spine muscle activation of walking on a narrow beam in virtual reality. *IEEE Transactions on Computer Graphics and Visualization*, in press.

[Armel and Ramachandran, 2003] Armel, K. and Ramachandran, V. (2003). Projecting sensations to external objects: evidence from skin conductance response. *Proceedings of the Royal Society, B, Biological Sciences*, 270(1523):1499–506.

[Avidan and Shashua, 1998] Avidan, S. and Shashua, A. (1998). Novel view synthesis by cascading trilinear tensors. *IEEE Transactions on Visualization and Computer Graphics*, 4(4):293–306.

[Ballard, 1981] Ballard, D. (1981). Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–121.

[Bayliss et al., 2004] Bayliss, J., Inverso, S., and Tentler, A. (2004). Changing the p300 brain computer interface. *Cyberpsychology & Behaviour*, 7(76):694–704.

[Bayliss, 2003] Bayliss, J. D. (2003). Use of the evoked potential p3 component for control in a virtual apartment. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 11(2):113–116. 1534-4320.

[Bente et al., 2003] Bente, G., Blens, H., and Krämer, N. C. (2003). Embodied conversational agents in e-commerce. do they sell? In *Proceedings of the 6th Annual Workshop on Presence*: http://www.temple.edu/ispr/prev_conferences/proceedings/2003/Bente,%20Blens,%20Kramer.pdf.

[Berti and Frassinetti, 2000] Berti, A. and Frassinetti, F. (2000). When far becomes near: Remapping of space by tool use. *Journal of Cognitive Neuroscience*, 12:415–420.

[Bertoncini et al., 2008] Bertoncini, M., Jacucci, G., Roveda, S., Tonguet, D., and Takala, T. (2008). Presence in performing digital art. In *The 11th Annual Workshop on Presence*: http://www.temple.edu/ispr/prev_conferences/proceedings/2008/bertoncini.pdf

[Biocca, 1997] Biocca, F. (1997). The cyborg's dilemma: Embodiment in virtual environments. *Journal of Computer-Mediated Communication*, 3(2):12–26.

[Biocca and Levy, 1995] Biocca, F. and Levy, M. R. (1995). *Communication in the age of virtual reality*. Publisher: L. Erlbaum Associates Inc., Hillsdale, NJ, USA.

[Birbaumer, 1997] Birbaumer, N. (1997). *Brain and Behavior: Past, Present and Future*, chapter Slow Cortical Potentials: their origin meaning and clinical use, pages 25–39. Tilburg University Press.

[Birbaumer et al., 2000] Birbaumer, N., Kubler, A., Ghanayim, N., Hinterberger, T., Perelmouter, J., Kaiser, J., Iversen, I., Kotchoubey, B., Neumann, N., and Flor, H. (2000). The thought translation device (ttd) for completely paralyzed patients. *IEEE Transactions on Rehabilitation Engineering*, 8(2):190–193.

[Black et al., 2008] Black, M., Friel, B., Jasak, J., Lombard, M., and Russ, A. (2008). An evaluation of portrayals of telepresence and romantic relationships in film and television. In *Proceedings of the 11th Annual Workshop on Presence*.

[Blake et al., 2007] Blake, E., Nunez, D., and Labuschagne, B. (2007). Longitudinal effects on presence: Suspension of disbelief or distrust of naive belief? In *Presence 2007*, pages 291–295.

[Blanchard and Epstein, 1978] Blanchard, E. B. and Epstein, L. H. (1978). *A Biofeedback Primer*. Publisher: Addison-Wesley.

[Blasi et al., 2008] Blasi, M. D., Ferraro, G., and Conti, F. (2008). The blog, the blogger and “the other” on the net. In *Proceedings of the 11th Annual Workshop on Presence*, pages 250–253.

[Botvinick, 2004] Botvinick, M. (2004). 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):756–756.

[Bouguila et al., 2004] Bouguila, L., Evequoz, F., Courant, M., and Hirsbrunner, B. (2004). Walking-pad: a step-in-place locomotion interface for virtual environments. In *ICMI ’04: Proceedings of the 6th international conference on Multimodal interfaces*, pages 77–81, New York, NY, USA. ACM.

[Bowman and Hodges, 1995] Bowman, D. A. and Hodges, L. F. (1995). User interface constraints for immersive virtual environment applications. Technical report, Atlanta: Graphics, Visualization and Usability.

[Bowman et al., 2001] Bowman, D. A., Johnson, D. B., and Hodges, L. F. (2001). Testbed evaluation of virtual environment interaction techniques. *Presence: Teleoperators & Virtual Environments*, 10(1):75–95.

[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 *VRAIS ’97: Proceedings of the 1997 Virtual Reality Annual International Symposium (VRAIS ’97)*, pages 45–52, Washington, DC, USA. IEEE Computer Society.

[Bowman et al., 2004] Bowman, D. A., Kruijff, E., LaViola, J. L., and Poupyrev, I. (2004). *3D User Interfaces: Theory and Practice*. Publisher: Addison-Wesley Professional.

[Bradski and Kaehler, 2008] Bradski, G. and Kaehler, A. (2008). *Learning OpenCV: Computer Vision with the OpenCV Library*. Publisher: O’Reilly Media.

[Brogni et al., 2003] Brogni, A., Slater, M., and Steed, A. (2003). More breaks less presence. In *Presence 2003*, pages 6–8, Aalborg, Denmark.

[Brogni et al., 2007] Brogni, A., Vinayagamoorthy, V., Steed, A., and Slater, M. (2007). Responses of participants during an immersive virtual environment experience. *The International Journal of Virtual Reality*, 6(2):1–10.

[Brooks, 1999] Brooks, F. P. (1999). What’s real about virtual reality? *IEEE Comput. Graph. Appl.*, 19(6):16–27.

[Cardillo et al., 2007] Cardillo, C., Russo, M., LeDuc, P., and Torch, W. (2007). *Quantitative EEG Changes Under Continuous Wakefulness and with Fatigue Countermeasures: Implications for Sustaining Aviator Performance*, volume 4565. Publisher: Springer Berlin/Heidelberg.

[Carmena et al., 2003] Carmena, J. M., Lebedev, M. A., Crist, R. E., O'Doherty, J. E., Santucci, D. M., Dimitrov, D. F., Patil, P. G., Henriquez, C. S., and Nicolelis, M. A. (2003). Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS Biol*, 1(2):193–208.

[Carrozzino et al., 2005] Carrozzino, M., Tecchia, F., Bacinelli, S., Cappelletti, C., and Bergamasco, M. (2005). Lowering the development time of multimodal interactive application: the real-life experience of the xvr project. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pages 270–273.

[Carvey et al., 2006] Carvey, A., Gouldstone, J., Vedurumudi, P., Whiton, A., and Ishii, H. (2006). Rubber shark as user interface. In *extended abstracts on Human Factors in Computing Systems (CHI '06)*, pages 634–639.

[Costa and McCrae, 1992] Costa, P.T.,Jr. & McCrae, R.R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual. Odessa, FL: Psychological Assessment Resources.

[Cruz-Neira et al., 1993] Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: the design and implementation of the cave. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 135–142, New York, NY, USA. ACM.

[Darken et al., 1997] Darken, R. P., Cockayne, W. R., and Carmein, D. (1997). The omni-directional treadmill: a locomotion device for virtual worlds. In *UIST '97: Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 213–221, New York, NY, USA. ACM.

[Dawson et al., 2000] Dawson, M. E., Schell, A. E., and Filion, D. L. (2000). *Handbook of psychophysiology*, chapter The electrodermal system, pages 200–223. Cambridge University Press.

[de Aguiar et al., 2005] de Aguiar, E., Theobalt, C., Magnor, M., and Seidel, H.-P. (2005). Reconstructing human shape and motion from multi-view video. In *2nd European Conference on Visual Media Production (CVMP)*, pages 42–49, London, UK.

[Donchin et al., 2000] Donchin, E., Spencer, K. M., and Wijesinghe, R. (2000). The mental prosthesis: assessing the speed of a p300-based brain-computer interface. *IEEE Trans Rehabil Eng*, 8(2):174–179.

[Dourish, 2001] Dourish, P. (2001). *Where the action is*. Publisher: MIT Press.

[Dow et al., 2007] Dow, S., Mehta, M., Harmon, E., MacIntyre, B., and Mateas, M. (2007). Presence and engagement in an interactive drama. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1475–1484, New York, NY, USA. ACM.

[Draper et al., 1998] Draper, J. V., Kaber, D. B., and Usher, J. M. (1998). Telepresence. *Human Factors*, 40(3):354–375.

[Duda and Hart, 1972] Duda, R. and Hart, P. (1972). Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15(1):11–15.

[Edlinger et al., 2009] Edlinger, G., Holzner, C., Groenegress, C., Guger, C., and Slater, M. (2009). Goal-oriented control with brain-computer interface. In *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, pages 732–740. Springer Berlin/Heidelberg.

[Edlinger et al., 2008] Edlinger, G., Krausz, G., Groenegress, C., Holzner, C., Guger, C., and Slater, M. (2008). Brain-computer interfaces for virtual environment control. In *13th International Conference on Biomedical Engineering*.

[Ehrsson et al., 2005a] Ehrsson, H. H., Holmes, N. P., and Passingham, R. E. (2005a). Touching a rubber hand: Feeling of body ownership is associated with activity in multisensory brain areas. *Journal of Neuroscience*, 25(45):10564–10573.

[Ehrsson et al., 2005b] Ehrsson, H. H., Kito, T., Sadato, N., Passingham, R. E., and Naito, E. (2005b). Neural substrate of body size: Illusory feeling of shrinking of the waist. *PLoS Biology*, 3(12), http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.0030412.

[Ellis, 1996] Ellis, S. (1996). Presence of mind: a reaction to thomas sheridan’s "further musings on the psychophysics of presence.". *Presence: Teleoperators and Virtual Environments*, 5(2):247–259.

[Elrod et al., 1992] Elrod, S., Bruce, R., Gold, R., Goldberg, D., Halasz, F., Janssen, W., Lee, D., McCall, K., Pedersen, E., Pier, K., Tang, J., and Welch, B. (1992). Liveboard: a large interactive display supporting group meetings, presentations, and remote collaboration. In *CHI ’92: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 599–607, New York, NY, USA. ACM.

[Engel et al., 2008] Engel, D., Curio, C., Tcheang, L., Mohler, B., and Bülthoff, H. H. (2008). A psychophysically calibrated controller for navigating through large environments in a limited free-walking space. In *VRST '08: Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pages 157–164, New York, NY, USA. ACM.

[Fairclough, 2008] Fairclough, S. (2008). Bci and physiological computings: Differences, similarities & intuitive control. In *Proceedings of CHI 08: Workshop on BCI and Computer Games*, http://hmi.ewi.utwente.nl/chi2008/chi2008_files/fairclough.pdf.

[Farwell and Donchin, 1988] Farwell, L. and Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6):510–523.

[Fitzmaurice et al., 1995] Fitzmaurice, G. W., Ishii, H., and Buxton, W. A. S. (1995). Bricks: laying the foundations for graspable user interfaces. In *CHI '95: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 442–449, New York, NY, USA. ACM Press/Addison-Wesley Publishing Co.

[Flach and Holden, 1998] Flach, J. and Holden, G. (1998). The reality of experience: Gibson's way. *Presence: Teleoperators and Virtual Environments*, 7:90–95.

[Foley et al., 1995] Foley, J. D., van Dam, A., Feiner, S. K., and Hughes, J. F. (1995). *Computer Graphics: Principles and Practice*. Publisher: Addison-Wesley Publishing.

[Freeman et al., 1999] Freeman, J., Avons, S. E., Pearson, D. E., and IJsselsteijn, W. A. (1999). Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence: Teleoper. Virtual Environ.*, 8(1):1–13.

[Friedman et al., 2007] Friedman, D., Leeb, R., Dikovsky, L., Reiner, M., Pfurtscheller, G., and Slater, M. (2007). Controlling a virtual body by thought in a highly-immersive virtual environment - a case study in using a brain-computer interface in a virtual-reality cave-like system. In *GRAPP*, pages 83–90, Barcelona, Spain.

[Fukunaga and Hostetler, 1975] Fukunaga, K. and Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *Information Theory, IEEE Transactions on*, 21(1):32–40.

[Gallagher and Cole, 1995] Gallagher, S. and Cole, J. (1995). Body image and body schema in a deafferented subject. *Journal of Mind and Behavior*, 16:369–90.

[Garau et al., 2008] Garau, M., Friedman, D., Ritter Weidenfeld, H., Antley, A., Brogni, A., and Slater, M. (2008). Temporal and spatial variations in presence: Qualitative analysis of interviews from an experiment on breaks in presence. *Presence - Teleoperators and Virtual Environments*, 17(3):293–309.

[Garau et al., 2005] Garau, M., Slater, M., Pertaub, D.-P., and Razzaque, S. (2005). The responses of people to virtual humans in an immersive virtual environment. *Presence - Teleoperators and Virtual Environments*, 14(1):104–116.

[Gibson, 1979] Gibson, J. J. (1979). *The ecological approach to visual perception*. Publisher: Houghton Mifflin, Boston, MA, United States.

[Gillies and Slater, 2005] Gillies, M. and Slater, M. (2005). Non-verbal communication for correlational characters. In *The 8th Annual International Workshop on Presence: Presence 2005*, pages 103–106.

[Grahe and Bernieri, 2002] Grahe, J. E. and Bernieri, F. J. (2002). Self-awareness of judgment policies of rapport. *Pers Soc Psychol Bull*, 28(10):1407–1418.

[Groenegress et al., 2009a] Groenegress, C., Slater, M., Holzner, C., and Guger, C. (2009a). Effects of bci use on presence in virtual reality. *submitted*.

[Groenegress et al., 2007] Groenegress, C., Slater, M., Tamke, M., and Thomsen, M. R. (2007). Spinoff - transferring energy between real and virtual worlds. In *Eurographics*, Prague.

[Groenegress et al., 2009b] Groenegress, C., Thomsen, M. R., and Slater, M. (2009b). Correlations between vocal input and visual response apparently enhance presence in virtual environments. *Cyberpsychology & Behaviour*, 12:in press.

[Guan et al., 2004] Guan, C., Thulasidas, M., and Wu, J. (2004). High performance p300 speller for brain-computer interface. In *IEEE International Workshop on Biomedical Circuits & Systems*, pages 13–16.

[Guger et al., 2008] Guger, C., Holzner, C., Groenegress, C., Edlinger, G., and Slater, M. (2008). Control of a smart home with a brain-computer interface. In *Proceedings of the 4th International Brain-Computer Interface Workshop 2008*, pages 339–342.

[Guger et al., 2009] Guger, C., Holzner, C., Groenegress, C., Edlinger, G., and Slater, M. (2009). Brain-computer interface for virtual reality control. In *Proceedings of ESANN 2009*, pages 443–448.

[Guger et al., 2001] Guger, C., Schlögl, A., Neuper, C., Walterspacher, D., Strein, T., and Pfurtscheller, G. (2001). Rapid prototyping of an eeg-based brain-computer interface (bci). *IEEE Trans. Rehab. Engng*, 9(1):49–58.

[Gysbers et al., 2004] Gysbers, A., Klimmt, C., Hartmann, T., Nosper, A., and Vorderer, P. (2004). Exploring the book problem: Text design, mental representations of space, and spatial presence in readers. In *Proceedings of the 7th International Workshop on Presence*, pages 13–20.

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

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

[Heldal et al., 2005a] Heldal, I., Steed, A., and Schroeder, R. (2005a). Evaluating collaboration in distributed virtual environments for a puzzle-solving task. In *HCI International 2005*, http://www.cs.ucl.ac.uk/staff/a.steed/heldal_hcii_cube.pdf.

[Heldal et al., 2005b] Heldal, I., Steed, A., Spante, M., Schroeder, R., Bengtsson, S., and Partanen, M. (2005b). Successes and failures in co-present situations. *Presence: Teleoperators & Virtual Environments*, 14(5):563–579.

[Huang et al., 2007] Huang, R.-S., Jung, T.-P., and Makeig, S. (2007). *Event-Related Brain Dynamics in Continuous Sustained-Attention Tasks*. Publisher: Springer Berlin/Heidelberg.

[Hwang and Lombard, 2006] Hwang, H. S. and Lombard, M. (2006). Understanding instant messaging: Gratifications and social presence. In *Proceedings of the 8th Annual Workshop on Presence*, pages 50–56.

[IJsselsteijn, 2002] IJsselsteijn, W. (2002). Elements of a multi-level theory of presence: Phenomenology, mental processing and neural correlates. In *Presence*, pages 245–259, Porto, Portugal.

[IJsselsteijn, 2005] IJsselsteijn, W. (2005). Towards a neuropsychological basis of presence. *Annual Review of CyberTherapy and Telemedicine*, 3:25–30.

[IJsselsteijn et al., 2000] IJsselsteijn, W., de Ridder, H., Freeman, J., and Avons, S. (2000). Presence: Concept, determinants and measurement. In *Proceedings of the SPIE, Human Vision and Electronic Imaging V*, volume 39, pages 3959–76, San Jose, CA.

[Imamizu et al., 2000] Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., Putz, B., Yoshioka, T., and Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*, 403(6766):153–4.

[Insko, 2001] Insko, B. E. (2001). Passive haptics significantly enhances virtual environments. Technical report, UNC.

[Ishii and Ullmer, 1997] Ishii, H. and Ullmer, B. (1997). Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI '97: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 234–241, Atlanta, Georgia, United States.

[Ishii et al., 1999] Ishii, H., Wisneski, C., Orbanes, J., Chun, B., and Paradiso, J. (1999). Pingpongplus: Design of an athletic-tangible interface for computer-supported cooperative play. In *CHI '99: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 394–401, Pittsburgh, Pennsylvania, USA.

[Iwata et al., 2005] Iwata, H., Yano, H., Fukushima, H., and Noma, H. (2005). Circulafloor. *IEEE Comput. Graph. Appl.*, 25(1):64–67.

[Jacob et al., 2008] Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., and Zigelbaum, J. (2008). Reality-based interaction: a framework for post-wimp interfaces. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 201–210, New York, NY, USA. ACM.

[Jia et al., 2007] Jia, C., Xu, H., and Hong, B. (2007). A human computer interface using ssvep-based bci technology. *Lecture Notes in Computer Science*, 4565:113–119.

[Kitamura et al., 2003] Kitamura, Y., Yamaguchi, Y., Imamizu, H., Kishino, F., and Kawato, M. (2003). Things happening in the brain while humans learn to use new tools. *Proceedings of the SIGCHI conference on Human factors in computing system*, 5(1):417–424.

[Krusienski et al., 2006] Krusienski, D., Sellers, E., Cabestaing, F., Bayoudh, S., McFarland, D., Vaughan, T., and Wolpaw, J. (2006). A comparison of classification techniques for the p300 speller. *Journal of Neural Engineering*, 3:299–305.

[Lakin and Chartrand, 2003] Lakin, J. L. and Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, 14:334–339.

[Lalor et al., 2005] Lalor, E. C., Kelly, S. P., Finucane, C., Burke, R., Smith, R., Reilly, R. B., and McDarby, G. (2005). Steady-state vep-based brain-computer interface control in an immersive 3d gaming environment. *EURASIP J. Appl. Signal Process.*, 2005(1):3156–3164.

[Lebedev and Nicolelis, 2006] Lebedev, M. A. and Nicolelis, M. A. (2006). Brain-machine interfaces: past, present and future. *Trends in Neurosciences*, 29(9):536 – 546.

[Lee and Nass, 2003] Lee, K. M. and Nass, C. (2003). Designing social presence of social actors in human computer interaction. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 289–296, New York, NY, USA. ACM.

[Leeb et al., 2004] Leeb, R., Kleinrath, C., Guger, C., Scherer, R., Friedman, D., Slater, M., and Pfurtscheller, G. (2004). Using a bci as a navigation tool in virtual environments. In *2nd International Brain-Computer Interface Workshop and Training*, pages 49–50.

[Leeb et al., 2007a] Leeb, R., Lee, F., Keinrath, C., Scherer, R., Bischof, H., and Pfurtscheller, G. (2007a). Brain-computer communication: Motivation, aim, and impact of exploring a virtual apartment. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 15(4):473–482.

[Leeb et al., 2008] Leeb, R., Lee, F., Kleinrath, C., Scherer, R., Bischof, H., and Pfurtscheller, G. (2008). Computer communication: Motivation, aim, and impact of exploring a virtual apartment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15(4):473–482.

[Leeb et al., 2005] Leeb, R., Scherer, R., Kleinrath, C., Guger, C., and Pfurtscheller, G. (2005). Exploring virtual environments with an eeg-based bci through motor imagery. *Biomedizinische Technik*, 52:86–91.

[Leeb et al., 2007b] Leeb, R., Settgast, V., Fellner, D., and Pfurtscheller, G. (2007b). Self-paced exploration of the austrian national library through thought. *INternational Journal of Bioelectromagnetism*, 9(4):237–244.

[Lessiter et al., 2001] Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: The itc-sense of presence inventory. *Presence: Teleoper. Virtual Environ.*, 10(3):282–297.

[Lok, 2004] Lok, B. (2004). Towards merging real and virtual spaces. *Communications of the ACM*, 47(8):48–53.

[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), online journal. http://jcmc.indiana.edu/vol3/issue2/lombard.html.

[Lympourides et al., 2009] Lympourides, V., Arvind, D., and Parker, M. (2009). Fully wireless, full body 3-d motion capture for improvisational performances. In *CHI ’09: Proceedings of the 27th international conference on Human factors in computing systems*, http://lister.cms.livjm.ac.uk/homepage/staff/cmsdengl/WBI2009/documents2009/CHI_2009_Paper_Edinburgh.pdf.

[Ma et al., 2007] Ma, Z., Gao, X., and Gao, S. (2007). Enhanced p300-based cursor movement control. *Lecture Notes in Computer Science*, 4565:120–126.

[MacDorman, 2005] MacDorman, K. F. (2005). Androids as an experimental apparatus: Why is there an uncanny valley and can we exploit it? In *Proceedings Of the CogSci 2005 Workshop: Toward Social Mechanisms of Android Science*, http://www.androidscience.com/proceedings2005/MacDormanCogSci2005AS.pdf.

[Maguire and Connaughton, 2006] Maguire, K. C. and Connaughton, S. L. (2006). Relational presence in distanced interdependent relationships. In *Proceedings of the 8th Annual Workshop on Presence*, page 35.

[Maravita and Iriki, 2004] Maravita, A. and Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences*, 8(2):79–86.

[McCloy and Stone, 2001] McCloy, R. and Stone, R. (2001). Science, medicine, and the future: Virtual reality in surgery. *BMJ*, 323(7318):912–915.

[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 Trans. Graph.*, 21(3):645–652.

[Meehan et al., 2003] Meehan, M., Razzaque, S., Whitton, M. C., and Brooks, Jr., F. P. (2003). Effect of latency on presence in stressful virtual environments. In *VR ’03: Proceedings of the IEEE Virtual Reality 2003*, pages 141–148, Washington, DC, USA.

[Middendorf et al., 2000] Middendorf, M., McMillan, G., Calhoun, G., and Jones, K. (2000). Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Transactions on Rehabilitation Engineering*, 8(2):211–4.

[Minsky, 1980] Minsky, M. (1980). Telepresence. *Omni*, pages 45–51.

[Nagai and Critchley, 2007] Nagai, Y. and Critchley, H. D. (2007). *New Research on Biofeedback*, chapter Novel Therapeutic Application of Galvanic Skin Response (GSR) Biofeedback to a Neurological Disorder: Mechanisms Underlying Biofeedback in Epilepsy Management, pages 1–31. Publisher: Nova Science Publishers.

[Nicolelis, 2001] Nicolelis, M. A. (2001). Actions from thoughts. *Nature*, 409(18):403–407.

[Nicolelis, 2003] Nicolelis, M. A. (2003). Brain–machine interfaces to restore motor function and probe neural circuits. *Nature Reviews Neuroscience*, 4:417–422.

[Nijholt et al., 2008] Nijholt, A., Tan, D., Pfurtscheller, G., Brunner, C., Millán, J. d. R., Allison, B., Graimann, B., Popescu, F., Blankertz, B., and Müller, K.-R. (2008). Brain-computer interfacing for intelligent systems. *IEEE Intelligent Systems*, 23(3):72–79.

[Noë, 2005] Noë, A. (2005). *Action in Perception*. Publisher: MIT Press.

[Norman, 1990] Norman, D. A. (1990). *The Design of Everyday Things*. Publisher: Doubleday.

[O’Regan et al., 2001] O’Regan, J., Myin, E., and Noë, A. (2001). Towards an analytic phenomenology: The concepts of bodiliness and grabbiness. In Carsetti, A., editor, *International Colloquium: Seeing and Thinking. Reflections on Kanizsa’s Studies in Visual Cognition*, pages 103–114, Rome.

[O’Regan and Noë, 2001] O’Regan, J. K. and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav Brain Sci*, 24(5):939–73.

[Oviatt, 2002] Oviatt, S. (2002). Multimodal interfaces. In Sears, J. J. and A., editors, *Handbook of Human-Computer Interaction*. Publisher: Lawrence Erlbaum, New Jersey.

[Pan et al., 2008] Pan, X., Gillies, M., and Slater, M. (2008). Male bodily responses during an interaction with a virtual woman. In *Intelligent Virtual Agents*, pages 89–96.

[Pan and Slater, 2007] Pan, X. and Slater, M. (2007). A preliminary study of shy males interacting with a virtual female. In *Proceedings of the 10th Annual Workshop on Presence*, pages 101–108.

[Peck et al., 2008] Peck, T., Whitton, M., and Fuchs, H. (2008). Evaluation of reorientation techniques for walking in large virtual environments. In *Transactions on Visualization and Computer Graphics*, pages 121–127.

[Perez-Marcos et al., 2009] Perez-Marcos, D., Slater, M., and Sanchez-Vives, M. V. (2009). Inducing a virtual hand ownership illusion through a brain-computer interface. *NeuroReport*, 20(6):589–594.

[Pertaub et al., 2002] Pertaub, D.-P., Slater, M., and Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators & Virtual Environments*, 11(1):68–78.

[Pfurtscheller, 1989] Pfurtscheller, G. (1989). Functional topography during sensorimotor activation studied with event-related desynchronization mapping. *Journal of clinical neurophysiology*, 6(1):75–84.

[Pfurtscheller et al., 2006] Pfurtscheller, G., Leeb, R., Kleinrath, C., Friedman, D., Neuper, C., Guger, C., and Slater, M. (2006). Walking from thought. *Brain Research*, 1071(1):145–152.

[Pfurtscheller and Lopes da Silva, 1999] Pfurtscheller, G. and Lopes da Silva, F. (1999). Event-related eeg/meg synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110:1842–1857.

[Pierce et al., 1997] Pierce, J. S., Forsberg, A. S., Conway, M. J., Hong, S., Zeleznik, R. C., and Mine, M. R. (1997). Image plane interaction techniques in 3d immersive environments. In *SI3D ’97: Proceedings of the 1997 symposium on Interactive 3D graphics*, pages 39–43, New York, NY, USA.

[Pinchbeck and Stevens, 2005] Pinchbeck, D. M. and Stevens, B. (2005). Presence, narrative and schemata. In *Proceedings of the 9th Annual Workshop of Presence*, pages 221–226.

[Poupyrev et al., 1997] Poupyrev, I., Weghorst, S., Billinghurst, M., and Ichikawa, T. (1997). A framework and testbed for studying manipulation techniques for immersive vr. In *Proceedings of VRST*, pages 21–28, Lausanne, Switzerland.

[Prothero et al., 1995] Prothero, J., Parker, D., Furness, T., and Wells, M. (1995). Towards a robust, quantitative measure for presence. Technical report p-95-8, HIT Lab.

[Razzaque et al., 2002] Razzaque, S., Swapp, D., Slater, M., Whitton, M. C., and Steed, A. (2002). Redirected walking in place. In *EGVE '02: Proceedings of the workshop on Virtual environments 2002*, pages 123–130, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

[Reynolds, 1987] Reynolds, C. (1987). Flocks, herds and schools: A distributed behavioral model. In *SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, volume 21, pages 25–34. ACM.

[Rheingold, 1991] Rheingold, H. (1991). *Virtual Reality*. Publisher: Summit Books, New York.

[Riva, 2005] Riva, G. (2005). Virtual reality in psychotherapy: Review. *Cyberpsychology & Behavior*, 8(3):220–230.

[Rizzolatti and Craighero, 2004] Rizzolatti, G. and Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27:169–192.

[Rizzolatti et al., 1997] Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (1997). Neuroscience: Enhanced: The space around us. *Science*, 277(5323):190–191.

[Rochat, 1998] Rochat, P. (1998). Self-perception and action in infancy. *Experimental Brain Research*, 123(1 - 2):102–109.

[Rose et al., 2005] Rose, F. D., Brooks, B. M., and Rizzo, A. A. (2005). Virtual reality in brain damage rehabilitation: Review. *Cyberpsychology & Behavior*, 8(3):241–262.

[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.

[Schlattmann et al., 2009] Schlattmann, M., Broekelschen, J., and Klein, R. (2009). Real-time bare-hands-tracking for 3d games. In *Proceedings of IADIS International Conference Game and Entertainment Technologies*, pages 59–66.

[Schlattmann et al., 2007] Schlattmann, M., Kahlesz, F., Sarlette, R., and Klein, R. (2007). Markerless 4 gestures 6 dof real-time visual tracking of the human hand with automatic initialization. *Computer Graphics Forum*, 26(3):467–476.

[Schroeder et al., 2001] Schroeder, R., Steed, A., Axelsson, A.-S., Heldal, I., Åsa Abelin, Wideström, J., Nilsson, A., and Slater, M. (2001). Collaborating in networked immersive spaces: as good as being there together? *Computers & Graphics*, 25(5):781 – 788.

[Schubert et al., 1999] Schubert, T., Friedmann, F., and Regenbrecht, H. (1999). Embodied presence in virtual environments. In Paton, R. and Neilson, I., editors, *Visual Representations and Interpretations*, pages 269–278. Springer, London.

[Schubert et al., 2001] Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–282.

[Sellers and Kübler, 2006] Sellers, E. W. and Kübler, A. (2006). Brain-computer interface research at the university of south florida cognitive psychophysiology laboratory: the p300 speller. *IEEE transactions on Neural Systems and Rehabilitation Engineering*, 14(2):221–4.

[Sheridan, 1992] Sheridan, T. (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments*, 1(1):120–126.

[Shneiderman and Plaisant, 1987] Shneiderman, B. and Plaisant, C. (1987). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Publisher: Addison-Wesley.

[Short et al., 1976] Short, J., Williams, E., and Christie, B. (1976). *The social psychology of telecommunications*. Publisher: John Wiley & Sons.

[Sivak and McKinley, 2009] Sivak, S. and McKinley, B. (2009). Using a full body interface to create a new way to play. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*. http://lister.cms.livjm.ac.uk/homepage/staff/cmsdengl/WBI2009/documents2009/SivakMcKinleyWholechi2009.doc.

[Slater, 1999] Slater, M. (1999). Measuring presence: A response to the witmer and singer questionnaire. *Presence: Teleoperators and Virtual Environments*, 8(5):560–566.

[Slater, 2002] Slater, M. (2002). Presence and the sixth sense. *Presence: Teleoperators and Virtual Environments*, 11(4):435–439.

[Slater, 2003] Slater, M. (2003). A note on presence terminology. *Presence-Connect*, 3, online journal: http://people.cs.uct.ac.za/ dnunez/reading/papers/presenceterminology.pdf.

[Slater, 2004] Slater, M. (2004). How colorful was your day? why questionnaires cannot assess presence in virtual environments. *Presence - Teleoperators and Virtual Environments*, 13(4):484–493.

[Slater, 2009] Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos Trans R Soc Lond*, in press.

[Slater et al., 2006a] Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., Pistrang, N., and Sanchez-Vives, M. V. (2006a). A virtual reprise of the stanley milgram obedience experiments. *PLoS ONE*, 1(1):http://www.plosone.org/article/fetchArticle.action?articleURI=info%3Adoi %2F10.1371%2Fjournal.pone.0000039.

[Slater et al., 2003] Slater, M., Brogni, A., and Steed, A. (2003). Physiological responses to breaks in presence: A pilot study. In *Presence 2003: The 6th Annual International Workshop on Presence*, pages 3–23.

[Slater and Garau, 2007] Slater, M. and Garau, M. (2007). The use of questionnaire data in presence studies: Do not seriously likert. *Presence: Teleoperators & Virtual Environments*, 16(4):447–456.

[Slater et al., 2006b] Slater, M., Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., Garau, M., Brogni, A., and Friedman, D. (2006b). Analysis of physiological responses to a social situation in an immersive virtual environment. *Presence - Teleoperators and Virtual Environments*, 15(5):553–569.

[Slater et al., 2009] Slater, M., Khanna, P., Mortensen, J., and Yu, I. (2009). Visual realism enhances realistic response in an immersive virtual environment. *IEEE Computer Graphics and Applications*, 29(3):76–84.

[Slater et al., 2008] Slater, M., Pérez Marcos, D., Ehrsson, H., and Sanchez-Vives, M. (2008). Towards a digital body: the virtual arm illusion. *Frontiers in Human Neuroscience*, 2(6):1–8.

[Slater and Steed, 2000] Slater, M. and Steed, A. (2000). A virtual presence counter. *Presence: Teleoperators and Virtual Environments*, 9(5):413–434.

[Slater et al., 1998] Slater, M., Steed, A., Maringelli, F., and McCarthy, J. (1998). The influence of body movement on subjective presence in virtual environments. *Human Factors*, 40(3):469–477.

[Slater et al., 1995a] Slater, M., Steed, A., and Usoh, M. (1995a). The virtual treadmill: a naturalistic metaphor for navigation in immersive virtual environments. In *VE '95: Selected papers of the Eurographics workshops on Virtual environments '95*, pages 135–148, London, UK. Springer-Verlag.

[Slater and Usoh, 1994] Slater, M. and Usoh, M. (1994). Body centred interaction in immersive virtual environments. In Magnenat Thalmann, N. and Thalmann, D., editors, *Artificial Life and Virtual Reality*, pages 125–148. John Wiley and Sons.

[Slater et al., 1995b] Slater, M., Usoh, M., and Steed, A. (1995b). Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Trans. Comput.-Hum. Interact.*, 2(3):201–219.

[Slater and Wilbur, 1997] Slater, M. and Wilbur, S. (1997). A framework for immersive virtual environments (five): speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 6(6):603–617.

[Sonka et al., 2007] Sonka, M., Hlavac, V., and Boyle, R. (2007). *Image Processing, Analysis, and Machine Vision*. Publisher: CL-Engineering.

[Spanlang, 2009] Spanlang, B. (2009). Halca a library for presence research. Technical report, Event Lab.

[Spanlang et al., 2008] Spanlang, B., Fröhlich, T., Antley, A., Fernandez Descalzo, V., Swapp, D., and Slater, M. (2008). Raving requires plausibility and immersion: A virtual fire experiment. In *Real Actions in Virtual Environments (RAVE)*.

[Steed et al., 2005] Steed, A., Brogni, A., and Vinayagamoorthy, V. (2005). Breaks in presence as usability criteria. In *HCI International 2005 - The 11th International Conference on Human Computer Interaction, Volume 9 - Advances in Virtual Environments Technology: Musings on Design, Evaluation, and Applications*, http://www.cs.ucl.ac.uk/staff/ucacajs/steed_hcii_bips.pdf.

[Steed, 2003] Steed, P. (2003). *Animating Real-Time Game Characters*. Publisher: Charles River Media.

[Stein and Meredith, 1993] Stein, B. E. and Meredith, M. A. (1993). *The Merging of the Senses*. Publisher: MIT Press.

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

[Sutherland, 1965] Sutherland, I. (1965). The ultimate display. In *Proceedings of IFIP Congress*, pages 506–508.

[Sutter, 1992] Sutter, E. (1992). The brain response interface: communication through visually-induced electrical brain responses. *Journal of Microcomputer Applications*, 15:31–45.

[Sutton et al., 1965] Sutton, S., Braren, M., Zubin, J., and John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, 150(3700):1187–1188.

[Swartout et al., 2006] Swartout, W., Gratch, J., Hill, R. W., Hovy, E., Marsella, S., Rickel, J., and Traum, D. (2006). Toward virtual humans. *AI Mag.*, 27(2):96–108.

[Synnott, 1993] Synnott, A. (1993). *The Body Social: Symbolism, Self, and Society*. Publisher: Routledge.

[Tarvainen et al., 2001] Tarvainen, M., Koistinen, A., Valkonen-Korhonen, M., Partanen, J., and Karjalainen, P. (2001). Analysis of galvanic skin responses with principal components and clustering techniques. *Biomedical Engineering, IEEE Transactions on*, 48(10):1071–1079.

[Tatai et al., 2003] Tatai, G., Laufer, L., Szaló, A., and Csordás, A. (2003). On transmitting emotions - experiments on multimodal presence over the internet. In *Proceedings of the 6th Annual Workshop on Presence*. http://www.temple.edu/ispr/prev_conferences/proceedings/2003/Tatai,%20Laufer,%20Szalo,%20Csordas.pdf.

[Taylor et al., 2001] Taylor, II, R. M., Hudson, T. C., Seeger, A., Weber, H., Juliano, J., and Helser, A. T. (2001). Vrpn: a device-independent, network-transparent vr peripheral system. In *VRST '01: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 55–61, New York, NY, USA. ACM.

[Theobalt et al., 2008] Theobalt, C., de Aguiar, E., Magnor, M., and Seidel, H.-P. (2008). *Three-dimensional television*, chapter Reconstructing human shape, motion and appearance from multi-view video, pages 29–58. Springer.

[Thornson et al., 2009] Thornson, C. A., Goldiez, B. F., and Le, H. (2009). Predicting presence: Constructing the tendency toward presence inventory. *Int. J. Hum.-Comput. Stud.*, 67(1):62–78.

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

[Turner and Turner, 2007] Turner, S. and Turner, P. (2007). Quality in qualitative methods. In *Proceedings of the 10th Annual Workshop on Presence*, pages 151–154.

[Underkoffler and Ishii, 1999] Underkoffler, J. and Ishii, H. (1999). Urp: A luminous-tangible workbench for urban planning and design. In *CHI ’99: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 386–393, Pittsburgh, Pennsylvania, USA.

.[Usoh et al., 1999] Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., and Brooks, Jr., F. P. (1999). Walking walking-in-place flying, in virtual environments. In *SIGGRAPH ’99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 359–364, New York, NY, USA. ACM Press/Addison-Wesley Publishing Co.

[Usoh et al., 2000] Usoh, M., Catena, E., Arman, S., and Slater, M. (2000). Using presence questionnaires in reality. *Presence - Teleoperators and Virtual Environments*, 9(5):497–503.

[van Dam, 1997] van Dam, A. (1997). Post-wimp user interfaces. *Communications of the ACM*, 40(2):63–67.

[van den Bos and Jeannerod, 2002] van den Bos, E. and Jeannerod, M. (2002). Sense of body and sense of action both contribute to self-recognition. *Cognition*, 85(2):177–87.

[Vanden Abeele et al., 2007] Vanden Abeele, M., Roe, K., and Pandelaere, M. (2007). Construct validation of the concepts social presence, emotional presence and connectedness and an application of zajonc's social facilitation theory to social presence research. In *Proceedings of the 10th Annual Workshop on Presence*, pages 215–224.

[Vaughan et al., 2006] Vaughan, T., McFarland, D., Schalk, G., Sarnacki, W., Krusienski, D., Sellers, E., and Wolpaw, J. (2006). The wadsworth bci research and development program: at home with bci. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(2):229–233.

[Vaughan et al., 1996] Vaughan, T., Wolpaw, J., and Donchin, E. (1996). Eeg-based communication: Prospects and problems. *IEEE Transactions on Rehabilitation Engineering*, 4(4):425–430.

[Venables and Christie, 1980] Venables, P. and Christie, M. (1980). *Techniques in Psychopsychology*, chapter Electrodermal activity, pages 2–67. John Wiler.

[Vidal, 1977] Vidal, J. J. (1977). Real-time detection of brain events in eeg. *Proceedings of the IEEE*, 65(5):633–641.

[Weinland et al., 2006] Weinland, D., Ronfard, R., and Boyer, E. (2006). Free viewpoint action recognition using motion history volumes. *Computer Vision and Image Understanding*, 104:249–257.

[Weiser, 1991] Weiser, M. (1991). The computer for the twenty-first century. *Scientific American*, 265(3):94–100.

[Weiser, 1993] Weiser, M. (1993). Ubiquitous computing. *Computer*, 26(10):71–72.

[Wellner, 1993] Wellner, P. (1993). Interacting with paper on the digitaldesk. *Communications of the ACM*, 36(7):87–96.

[Whitton et al., 2005] Whitton, M., Lok, B., Insko, B., and Brooks, F. (2005). Integrating real and virtual objects. In *International Conference on Virtual Reality 2005*,
http://www.cs.unc.edu/ whitton/ExtendedCV/Papers/2005-HCII-Whitton-MIxedEnvs.pdf.

[Wiederhold and Wiederhold, 2009] Wiederhold, B. K. and Wiederhold, M. D. (2009). *Virtual Reality Therapy for Anxiety Disorders: Advances in Evaluation and Treatment*. Publisher: American Psychological Association.

[Witmer and Singer, 1998] Witmer, B. and Singer, M. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240.

[Wolpaw et al., 2002] Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113:767–791.

[Zahorik and Jenison, 1998] Zahorik, P. and Jenison, R. L. (1998). Presence as being-in-the-world. *Presence: Teleoperators and Virtual Environments*, 7(1):78–89.

[Zeltzer, 1992] Zeltzer, D. (1992). Autonomy, interaction, and presence. *Presence - Teleoperators and Virtual Environments*, 1(1):127–132.

[Zhao et al., 2009] Zhao, Q., Zhang, L., and Cichocki, A. (2009). Eeg-based asynchronous bci control of a car in 3d virtual reality environments. *Chinese Science Bulletin*, 54(1):78–87.