Being There Together: Experiments on Presence in Virtual Environments (1990s)

This collection of papers represents work during the 1990s on the concept of presence in virtual environments. I put this collection together as notes for a short lecture series that I did at the Swiss Federal Institute of Science and Technology (EPFL), invited by Professor Daniel Thalmann, entitled 'Towards a Science of Virtual Reality' in September 1998.

The earlier research was carried out in the Department of Computer Science, Queen Mary College, University of London, and then from 1996 at the Department of Computer Science, University College London. Each Chapter is a pre-publication copy of a paper that was published, as in the references below.

Chapter 1 *The Role of Presence in Virtual Environments* [1]
Chapter 2 *Representations Systems and the Virtual Body* [2]
Chapter 3 *Depth of Presence in a Fairy Tale Setting* [3]
Chapter 4 *The Power of Shadows* [4]
Chapter 5 *Performance in a 3D Chess World* [5]
Chapter 6 *Body Centred Interaction* [6]
Chapter 7 *Walking to the Precipice* [7]
Chapter 8 *Body Movement in the Garden* [8]
Chapter 9 *Counting Presence in the 3D Chess World* [9]
Chapter 10 *Being Together Through Virtual Touch* [10,11]
Chapter 11 *Small Group Meetings* [12,13,14].

The most recent paper in the collection above was published in 2000. Since then new ideas about presence have been presented which supersede the earlier ones, that were encapsulated in the review [15]. Nevertheless the new ideas clearly have their roots in those earlier ones. For example, we placed great emphasis on the importance of making use of whole body movements for perception in virtual reality (Chapters 6-8), and this figures prominently in the new theory, based on the idea of sensorimotor contingencies. The new approach can be found in [16], and with ideas about measurement in [17]. Regarding the discussion of the influence of shadows (Chapter 4) there have been recent developments discussed in [18,19].

The virtual body (Chapter 2) has received a lot of attention in recent years. At the time we first investigated this we did not realise the significance of the fact that participants in a virtual reality would have the illusion that the virtual body that they saw (however crude) somehow became 'their body'. The use of virtual reality in how the brain represents the body is now a major topic in cognitive neuroscience [20,21,22,23,24,25,26,27,28].

Mel Slater, Barcelona, February 2013.

References

1. Slater M, 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: 603-616.
2. Slater M, Usoh M (1994) Representation Systems, Perceptual Position and Presence in Virtual Environments. Presence-Teleoperators and Virtual Environments 2: 221-234.
3. Slater M, Usoh M, Steed A (1994) Depth of Presence in Immersive Virtual Environments. Presence-Teleoperators and Virtual Environments 3: 130-144.
4. Slater M, Usoh M, Chrysanthou Y (1995) The influence of dynamic shadows on presence in immersive virtual environments. Selected papers of the Eurographics workshops on Virtual environments '95. Barcelona, Spain: Springer-Verlag. pp. 8-21.
5. Slater M, Linakis V, Usoh M, Kooper R (1996) Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess. ACM virtual reality software and technology (VRST) 163-172.
6. Slater M, Usoh M (1994) Body Centred Interaction in Immersive Virtual Environments. In: Thalmann NM, Thalmann D, editors. Artificial Life and Virtual Reality: John Wiley and Sons. pp. 125-148.
7. Slater M, Usoh M, Steed A (1995) Taking steps: the influence of a walking technique on presence in virtual reality. ACM Trans Comput-Hum Interact 2: 201-219.
8. Slater M, Steed A, McCarthy J, Maringelli F (1998) The influence of body movement on subjective presence in virtual environments. Human Factors 40: 469-477.
9. Slater M, Steed A (2000) A virtual presence counter. Presence-Teleoperators and Virtual Environments 9: 413-434.
10. Ho C, Basdogan C, Slater M, Durlach N, Srinivasan M. An experiment on the influence of haptic communication on the sense of being together; 1998. Citeseer.
11. Basdogan C, Ho C-H, Srinivasan MA, Slater M (2000) An experimental study on the role of touch in shared virtual environments. ACM Trans Comput-Hum Interact 7: 443-460.
12. Slater M, Sadagic A, Usoh M, Schroeder R (2000) Small-group behavior in a virtual and real environment: A comparative study. PRESENCE -Teleoperators and Virtual Environments 9: 37-51.
13. Steed A, Slater M, Sadagic A, Bullock A, Tromp J (1999) Leadership and collaboration in shared virtual environments. Virtual Reality, 1999 Proceedings, IEEE: 112-115.
14. Tromp J, Bullock A, Steed A, Sadagic A, Slater M, et al. (1998) Small group behavior experiments in the Coven project. IEEE Computer Graphics and Applications 18: 53-63.
15. Sanchez-Vives MV, Slater M (2005) From Presence to Consciousness through Virtual Reality. Nature Reviews Neuroscience 6: 332-339.
16. Slater M (2009) Place Illusion and Plausibility can lead to realistic behaviour in immersive virtual environments. Philos Trans R Soc Lond 364: 3549-3557.
17. Slater M, Spanlang B, Corominas D (2010) Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. ACM Transactions on Graphics 29: Paper: 92.
18. Slater M, Khanna P, Mortensen J, Yu I (2009) Visual Realism Enhances Realistic Response in an Immersive Virtual Environment. IEEE Computer Graphics and Applications 29: 76-84.

19. Yu I, Mortensen J, Khanna P, Spanlang B, Slater M (2012) Visual realism enhances realistic response in an immersive virtual environment - Part 2. IEEE Computer Graphics and Applications 32: 36-45.
20. Slater M, Perez-Marcos D, Ehrsson HH, Sanchez-Vives M (2008) Towards a digital body: The virtual arm illusion. Front Hum Neurosci 2: doi:10.3389/neuro. 3309.3006.2008.
21. Ehrsson HH (2007) The experimental induction of out-of-body experiences. Science 317: 1048-1048.
22. Petkova VI, Ehrsson HH (2008) If I Were You: Perceptual Illusion of Body Swapping. PLoS ONE 3(12): e3832. doi:10.1371/journal.pone.0003832.
23. Lenggenhager B, Tadi T, Metzinger T, Blanke O (2007) Video ergo sum: Manipulating bodily self-consciousness. Science 317: 1096-1099.
24. Slater M, Spanlang B, Sanchez-Vives M, Blanke O (2010) First person experience of body transfer in virtual reality. PLoS ONE: e10564. doi:10510.11371/journal.pone. 0010564.
25. Normand JM, Giannopoulos E, Spanlang B, Slater M (2011) Multisensory Stimulation Can Induce an Illusion of Larger Belly Size in Immersive Virtual Reality. PLoS ONE 6: e16128. doi:16110.11371/journal.pone.0016128.
26. Kilteni K, Groten R, Slater M (2012) The Sense of Embodiment in Virtual Reality. Presence: Teleoperators and Virtual Environments 21: 373-387.
27. Kilteni K, Normand J-M, Sanchez-Vives MV, Slater M (2012) Extending Body Space in Immersive Virtual Reality: A Very Long Arm Illusion. PLoS ONE 7: e40867.
28. Yuan Y, Steed A (2010) Is the Rubber Hand Illusion Induced by Immersive Virtual Reality? IEEE Virtual Reality. Waltham, MA, USA. pp. 95 - 102.

Being There Together

Experiments on Presence in Virtual Environments

Mel Slater

Anthony Steed

Martin Usoh

Department of Computer Science
University College London.

Draft 1, 12th September, 1998, MS.

Revision History:

CHAPTER 11 *Small Group Meetings* *189*

CHAPTER 12 *References* *209*

Preface

In the summer of 1990 I went to the annual SIG-GRAPH conference held that year in Dallas Texas. I had the opportunity to visit the VPL stand on the exhibition floor, in order to try out the 'virtual reality' equipment. The term 'virtual reality' had been recently coined by Jaron Lanier, and although many of the ideas had been around for several years, in fact dating back to Ivan Sutherland's concepts in the 1960s, it was of late causing huge interest both in the technical world of computer science, and in the media and wider public.

There was a very long queue throughout the exhibition period at the VPL stand, but I somehow managed to avoid this, having recently spoke on the same platform in London as one of the VPL executives.

I sat down ready to enter the virtual reality, an assistant helped me with the data glove, put the head-mounted display on me, adjusted the ear- and eye-phones, and I was 'there'. Where? Looking at some rather fat pixels, and feeling disappointed. The assistant said: 'You can move your head and look around.' I turned my head, and immediately I was somewhere else. I was in a room, and floating across the floor. I could see my disembodied 'hand'. I drifted towards a window, because 'outside' I could hear music. I looked through the window, looking out over a cliff, and down below was a boat. The music was coming from the boat.

I floated slowly down towards the boat, and then started looking around for the source of the music. The

boat seemed deserted though. Then an unexpected voice from far away said: 'YOUR TIME IS UP!' Hands were unscrewing the head-mounted display, it was taken off, and disoriented I was back in the rich, amazingly bright and colourful environment that we call 'reality'.

This experience 'changed my life' as they say. I realised that this was not just another way to view interactive 3D graphics, but a profoundly different media, a different coupling between people and computers, a way of placing a human inside a computer generated environment. Although the underlying graphics technology was no different to what is used for display on a single screen, the relationship between the 'user' and the portrayed environment is completely different. The 'user' is part of the environment, not looking at it, or interacting with it from the outside. Why 'user'? User seemed a strange word in this context - the concept of 'participant' was much more appropriate - people are 'in' these virtual environments, not simply 'using' them.

This notion of being in the virtual environment was embodied in a concept called 'presence'. Presence seemed to be the major distinguishing feature, the new property that 'virtual reality' offered. Studying what contributes to this sense of presence offered an approach to constructing a theory of virtual reality, to understanding what happens when people enter into virtual environments. There had been substantial and extremely successful research for more than 20 years on human factors and user-interface design with respect to 2D human-computer interaction. The outcome of this research was a major factor in the revolution that led to

a computer in almost every home, computers usable by people. However, this new type of interaction, needed a new approach. Several possibilities existed, such as, for example, understanding how people perform in virtual environments, how training in virtual environments relates to their performance in the real world, understanding human perception and motor performance in the context of virtual environments, taking a traditional human factors point of view. The question of presence seemed to me to be far more interesting, and unusual, and I decided to go down this road.

During 1991 and 1992 I was most of the time at the University of California Berkeley, teaching and researching in computer graphics. Nevertheless in January 1992 a new project started in London, of which I was the principal investigator, in the general area of Virtual Reality for Architecture. This was in conjunction with the newly formed London Parallel Applications Centre at Queen Mary and Westfield College, University of London, where I was based at the time, and with Thorn EMI and Division Ltd. A research assistant, Dr. Martin Usoh was recruited for the project, and he started work in January 1992 - with supervision, electronically from afar by myself.

The first system that we had was a Division System, running on a multi-processor transputer based machine, and software called dVS. Martin Usoh spent several months setting up this software, hardware and becoming familiar with this.

When I returned to London in the summer of 1992, the system was fully operational, and Martin was well-

acquainted with the software. (In those days we had source-code, which made the learning path somewhat easier).

My first experience with the Division System was again interesting. I was walking around in a virtual corridor, feeling completely there (though rather sick with so-called 'simulator sickness). My hand was represented as an arrow which responded to hand and wrist movements. At one point I looked down at my body. I was extremely surprised to find that it was *not there!* This immediately 'threw me out' of the presence experience.

My interest in presence had been theoretical up to that point. I have a strong background in statistical methods and experimental design, and I decided that some traditional experimentation was required to understand empirically what was going on. Working mostly from intuition, an experiment was designed in order to test the hypothesis that presence would, on the average, be stronger for people if they had a virtual body, compared to those who only had a disembodied virtual hand.

Martin designed and implemented the virtual body, and the experiment was carried out. This led us on a path that has ultimately resulted in this book - many hundreds of experimental subjects later.

In October 1992 Anthony Steed became my PhD student at QMW. He also became very interested in the experimental approach to understanding virtual reality, and quickly became an expert in the dVS software. His own PhD topic was again on the theme of building tools

for this new media that were native to the media itself rather than copied from existing (inevitably 2D) ideas. Martin was by this time employed on a project that was looking at geometric design in the context of VR. Anthony constructed a system that would allow changes to interactive behaviours in the VR, but changes carried out by a participant. In other words the functionality of how the virtual environment operated could be changed as it were by 'super participants' (not users!) from inside the VR.

In November 1995 what had been the virtual environments and graphics group at QMW moved with me to University College London. Martin moved immediately, and Anthony after he had finished his PhD in 1996. A new project, funded by the European Union under the ACTS framework had started, called 'Collaborative Virtual Environments'. This was in conjunction with many partners in the UK and in the rest of Europe. Over the next three years at UCL we carried out many experiments on presence in virtual environments, and then broadened this to understanding what happens when people meet in a virtual environment. The important issue seemed to be: how would a virtual environment meeting compare to the same people meeting in a real environment? How do the everyday attitudes, beliefs and behaviours carry over to a virtual environment. Are shy people still shy? Do 'leaders' remain leaders? A series of studies was embarked on, in conjunction with the sociologist Ralph Schroeder (now at Chalmers University in Sweden) and Amela Sadagic (a PhD student at UCL) in order to study these questions. These issues

were also taken up in two international trials as part of the COVEN project.

During February to April 1998 I was Visiting Scientist at MIT's Research Laboratory for Electronics. My host was Senior Scientist Nat Durlach, in the RLE Sensory Communications Group. With a background in computer graphics, my main focus had been on the visual aspects of virtual reality. Over a dinner conversation one night, Nat raised the idea of the potential importance of touch feedback for enhancing the sense of 'togetherness' between people sharing the same visual VE. The RLE has a 'Touch Laboratory' led by Dr. M. A. Srinivasan, where there is significant research on the Phantom haptic device. I proposed an experiment that would explore this sense of 'being together' with two remote people carrying out a task together, but feeling force feedback from one another. This was refined into an actual experiment together with Nat Durlach and Dr. C. Basdogan, and the scenario was implemented by C. Ho. This forms the basis for Chapter 10, which reports the first ever study on the relationship between haptic feedback and the sense of 'being together', as well as the influence of haptic feedback on successful performance of the task.

Apart from Chapter 10, all the studies reported in this book are from experiments carried out in the University of London (at QMW from 1992 until 1995, and thereafter at UCL).

I decided to try to put these together as a book due to an invite by Professor Daniel Thalmann at EPFL in Switzerland, to give a series of six lectures on a topic of my choice to the EPFL students. This seemed to be a good

opportunity to bring several years' work together. Some of this will no doubt seem dated. Also this is not a 'text book' in the sense that there is no deep study of the contributions of other authors. Rather we put forward the specific methodology developed over the past few years in London, and some of the interesting results that have been obtained.

Mel Slater,

London, September, 1998.

CHAPTER 1

The Role of Presence in Virtual Environments

"We modern, civilised, indoors adults are so accustomed to looking at a page or a picture, or through a window, that we often lose the feeling of being *surrounded* by the environment, our sense of the *ambient* array of light... We live boxed up lives." (Gibson, 1986)

Summary

This chapter reviews the concepts of immersion and presence in virtual environments. We propose that the degree of immersion can be objectively assessed as the characteristics of a technology, and has dimensions such as the extent to which a display system can deliver an inclusive, extensive, surrounding and vivid illusion of virtual environment to a participant. Other dimensions of immersion are concerned with the extent of body matching, and the extent to which there is a self-contained plot in which the participant can act and in which there is an autonomous response. Presence is a state of consciousness that may be concomitant with immersion, and is related to a sense of being in a place. Presence governs aspects of autonomic responses and more gross behaviour of a participant in a VE. The chapter considers single and multi-participant shared environments.

1. Introduction: Through the Looking Glass

Those of us old enough will remember working in institutions many years ago that had a special "computer room". This was a glass encased temperature controlled room, with banks of large whirling tape drives, discs,

large blue boxes with lots of flashing lights, attended by priest-like operators in white coats. Day after day we would pass by that room, and maybe we were able to see through the glass, to observe that essentially sacred place and the objects of worship and rites and rituals within it.

One of the authors had quite an unusual experience one day after about four years of passing by such a Computer Room in College: he had to go inside it. It was rather a shock. What had been seen on the outside, only ever through the glass, only ever from the limited range of viewpoints afforded by the architecture and room layout, was now suddenly surrounding - he was inside it, he saw (and experienced) the computer room in a way that had never been possible before for him, in a way that was impossible from the outside.

When we look at a TV screen or movie, it is much the same as looking through this glass - except that the scenario and unfolding events are typically distant in place and time. The glass of the TV screen forms a discontinuity between the place of our current reality, and the reality showing through the display. This discontinuity between different spatial and temporal realities, and its sudden unexpected collapse, is a recurring theme in popular culture. Considering this in relation to a Robert Henlein novel (The Unpleasant Profession of Jonathon Hoag), regarding a scene where a couple in a car roll down a window pane to find an Absolute Nothingness outside, Slavoj Z_ iz _ek (1991) writes:

> '...To those sitting inside a car, outside reality appears slightly distant, the other side of a barrier or screen materialised by the glass. We perceive external reality, the world outside the car, as "another reality", another mode of reality, not immediately continuous with the reality inside the car. The proof of this discontinuity is the uneasy feeling that overwhelms us when we suddenly roll down the window-pane and allow external reality to strike us with the proximity of its material presence. Our uneasiness consists in the sudden experience of how close really is what the windowpane, serving as a kind of protective screen, kept at a safe distance. But when we are safely inside the car, behind the closed windows, the external objects are, so to speak, transposed into another mode. They appear to be fundamentally "unreal", as if their reality has been suspended, put in parenthesis - in short, they appear as a kind of cinematic reality projected onto the screen of the windowpane. It is precisely this phenomenological experience of the barrier separating inside from outside, this feeling that the outside is ultimately "fictional", that produces the horrifying effect of the final scene in Henlein's novel. It is as if, for a moment, the "projection" of the outside reality had stopped working, as if, for a moment, we had been confronted with the formless grey, with the emptiness of the screen...'

When we look at a computer screen the scenario and events are now not "real" but computer generated: the environment that we are looking at is "virtual", it is a representation of something - some underlying process, or computation, rather than what it appears to be.

The grand aim of immersive virtual environments research is to be able to realise that same "stepping through the glass" or "rolling down the window" with respect to computer generated environments, as can be experienced when stepping through a barrier that in normal circumstances screens some aspect of reality from us. But this stepping through the barrier has some paradoxical elements: on the one hand, it is a surprise, when the previously remote suddenly becomes immediate, it is essentially *unreal*. Simultaneously though, we wish to preserve something in the passage through the barrier, that is the sense of our *self* being in a place, the sense that we are really through the barrier - that is, preserving the invariance of our sense of "being there", commonly referred to as the

sense of *presence*, or tele-presence. As has been argued by Steur (1992) presence is the central goal of "virtual reality", perhaps a defining feature.

The need to maintain a sense of presence even after passing through the barrier therefore has become a guiding principle for our research. In this chapter therefore we will review our approach to the definition of presence, and the emerging model for understanding the factors that influence this. We will also briefly consider the concept of presence in shared environments, and then return to how this can be a guide for research.

2. Immersion and Presence

2.1 Immersion

We distinguish between immersion and presence. *Immersion* is a description of a technology, and describes the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant. *Inclusive (I)* indicates the extent to which physical reality is shut out. *Extensive (E)* indicates the range of sensory modalities accommodated. *Surrounding (S)* indicates the extent to which this virtual reality is panoramic rather than limited to a narrow field. *Vivid* (V) indicates the resolution, fidelity, and variety of energy simulated within a particular modality (for example, the visual and colour resolution). Vividness is concerned with the richness, information content, resolution and quality of the displays.

These aspects of immersion are concerned with display of information. *Matching* requires that there is match between the participant's proprioceptive feedback about body movements, and the information generated on the displays. A turn of the head should result in a corresponding change to the visual display, and, for example, to the auditory displays so that sound direction is invariant to the orientation of the head. Matching requires body tracking, at least head tracking, but generally the greater the degree of body mapping, the greater the extent to which the movements of the body can be accurately reproduced.

Immersion requires a self-representation in the VE - a Virtual Body (VB). The VB is both part of the perceived environment, and represents the being that is doing the perceiving. Perception in the VE is centred on the position in virtual space of the VB - e.g., visual perception from the viewpoint of the eyes in the head of the VB (*egocentric* as opposed to *exocentric*, Ellis, 1991).

Each of these dimensions of immersion has, in principle, associated scales, indicating the extent of their realisation. For example, "surrounding" can be delivered by a small external screen at one extreme and a wide field of view HMD, or a CAVE system at the other. "Inclusive" in the ideal situation would, for example, have the HMD completely weightless, so that this aspect of external reality is not perceived by the participant. "Vivid" would include, for example, the quality of the visual rendering (from wire frame to photo-realism) as well as more basic considerations such as the pixel resolution.

Each of these dimensions exists on multiple levels. The most fundamental levels may correlate with the responses of the autonomic nervous system - for example, whether the VE visual display has the capability to induce changes in visual accommodation and vergence (Ellis, 1991). Higher levels may correlate with cognitive responses and behaviours. For example, whether or not the system can exhibit dynamically changing shadows may influence a participant's behaviour in certain tasks such as picking up objects, or aiming projectiles before firing them (Slater, Usoh, Chrysanthou, 1995).

The case of "matching" requires at the most basic level a minimal lag between motor actions and the corresponding system response. At a higher level matching has implications for the interaction paradigms employed. The concept of "body centred interaction" (Slater and Usoh, 1994), developed as a result of these ideas, requires that actions be carried out in a way that maximises the match between proprioception and sensory feedback at the perceptual and cognitive level. A very straightforward example, is that ideally a participant should virtually walk by really walking - in which case the whole body movements associated with walking match the corresponding optical flow.

Finally we mention *plot*. This is the extent to which the VE in a particular context presents a story-line that is self-contained, has its own dynamic, and presents an alternate unfolding sequence of events, quite distinct from those currently going on in the "real world". This includes Zeltzer's (1992) notion of "autonomy" (the extent to which objects in the VE have their own independent behaviour) and also the response of other virtual actors to actions of participants (Heeter, 1992). It also includes Zeltzer's notion of "interaction", that is the extent to which the participant can influence the unfolding of events, and effect changes to the virtual world. Plot is in a sense the extent to which the VE can potentially "remove" the participant from everyday reality and realise and act in an alternative self-contained world with its own drama in which the individual can participate.

2.2 Presence

Immersion can be an objective and quantifiable description of what any particular system does provide. Presence is a state of consciousness, the (psychological) sense of being in the virtual environment. Presence has been studied by many researchers in recent years, for example (Heeter, 1992; Held and Durlach, 1992; Loomis, 1992; Sheridan, 1992; Steur, 1992; Barfield and Weghorst, 1993; Barfield et. al., 1995). The fundamental idea is that participants who are highly present should experience the VE as more the engaging reality than the surrounding physical world, and consider the environment specified by the displays as places visited rather than as images seen. Behaviours in the VE should be consistent with behaviours that would have occurred in everyday reality in similar circumstances. Presence therefore requires that the participant identify with the VB - that its movements are his/her movements, and that the VB comes to "be" the body of that person in the VE.

There are several working hypotheses that have emerged from and latterly guided a number of our practical experiments:

(a) Presence is both a subjective and objective description of a person's state with respect to an environment. The subjective relates to their evaluation of their degree of "being there", the extent to which they think of the virtual

environment as "place like" (subject to suspension of disbelief). The objective is an observable behavioural phenomenon, the extent to which individuals behave in a VE similar to the way they would behave in similar circumstances in everyday reality. The subjective may be correlated with the higher levels of immersion mentioned above. The objective may be correlated with more fundamental aspects of immersion.

(b) We think of presence as an increasing function of immersion in all its aspects. However, the impact of the display aspects (I, S, E, V) is mediated through two filters - the application or task context and the perceptual requirements of the individual. The first is obvious - for example, an application concerned with understanding the relationship between location within a chamber and the auditory quality of an orchestra must have high quality auditory rendering to be meaningful, whereas the visual representation is less important. Secondly, individuals seem to differ in their preference for information in the various modalities to enable a successful construction of their internal world models. For one person the absence of auditory information might be a crucial hindrance, whereas for another it might be hardly noticeable.

(c) The more the "plot" line potentially removes a person from everyday reality, and presents an alternate self-contained world, the greater the chance for presence. On the subjective side the more that a person is susceptible to displacement of their sense of reality, the greater the chance for presence. This might be measured, for example, by their degree of susceptibility to hypnosis.

2.3 Influence of Immersion on Presence

In their 1992 paper Held and Durlach (op. cit.) note regarding understanding of the factors that explain presence that "there is no scientific body of data and/or theory delineating the factors that underlie the phenomenon". Although this remains largely true, there have since been a few experimental studies which we now briefly consider in relation to some of the aspects of immersion considered above.

(a) Inclusive

Held and Durlach argue that presence requires that the displays be free from signals that indicate the existence of the device, which, of course, belongs to the physical rather than the virtual reality. Such signals would include three categories - those directly due to the information display systems, such as aliases and slow update rates; the input systems - such as interference caused by metallic objects in the electro-magnetic sensors; and the physical properties of the devices themselves - weight, cables, and so on. In our first experimental study (Slater and Usoh, 1992) we found from questionnaire responses after an experiment that in answer to the open question "Were there any circumstances that especially *decreased* your sense of being 'really there'?" 4 out of 17 subjects mentioned outside events including the voice of the experimenter, and 6/17 mentioned poor screen updates, low resolution, and high lag. However, when in the same study a deliberate attempt was made to cause outside interference (making a loud and incongruous noise by dropping a cup and saucer) those who reported the highest sense of presence actually *incorporated* this noisy event into their VE experience - i.e., the source was experienced as if it had occurred from within the environment rather than from external reality. (This recalls Freud's observations in the *Interpretation of Dreams*, that dreamers weave outside events into the fabric of their dreams.

He mentions Maury's famous dream about being guillotined as being prompted by something falling on his neck while sleeping).

In a small pilot study to study the effects of auditory phenomena on presence Patel (1994) carried out an experiment where the subjects were grouped according to the quality of sound they received - sound only from the real world of the laboratory, white noise generated by the HMD speakers, non-directional sound generated by the speakers, and finally spatialised directional sound. The result was that the largest change in the influence on presence was from the "no virtual sound" condition to the "white noise" condition - suggesting that the white noise isolated the subject from the real world sounds, supporting this notion of inclusion.

Finally, a study by Barfield and Hendrix (1995) examined the influence on reported presence of update rate. They found that there was such an influence, that presence generally increased with increasing update rate, but that the reported presence was approximately constant between about 15Hz and 20Hz.

(b) Vividness

Welch et. al (1996) reported an experiment with a driving simulator where two levels of pictorial realism were presented. There was a significant difference in level of reported presence between the two levels of pictorial realism, with the more realistic resulting in a higher level of reported presence. Hendrix and Barfield (1996a) studied the effects of stereopsis, and geometric field of view on subjective presence. Each of these significantly affected reported presence, with stereopsis and a wider geometric field of view each positively correlated with the presence score.

We mentioned shadows as an example of "high level" vividness. In the cited study subjects were asked to carry out a task involving the selection and firing of a projectile at a target. The extent of dynamic shadows was an independently varied factor, and all subjects carried out the same task. In this experiment presence was measured subjectively, using a questionnaire, but also there was an attempt to measure "behavioural presence" - in this case the discrepancy of a pointing angle between a real and virtual source (the greater the angle the more that the subject was influenced by the virtual). The extent to which the subjects experienced dynamic shadows was positively and significantly correlated with both subjective and behavioural scales of presence.

In a recent study Uno and Slater (1997) examined the influence of the visual simulation of the physical laws on reported presence. In this study with 18 subjects, each was exposed to differing combinations of elasticity, friction, and collision response in the context of a virtual bowling alley. It was found that in this application, the more realistic simulation of friction was significantly and positively associated with reported presence, but that more accurate simulations of elasticity and collision response did not have such an effect.

(c) Proprioceptive Matching

In the same study by Welch et. al., delay in visual feedback was another independent factor. A higher level of presence was reported under the condition of minimal delay, and this was a more important factor than the level of pictorial realism. Hendrix and Barfield (1996a) found that head-tracking significantly increased the reported

sense of presence in an experimental study, and also led to subjects becoming more animated in the use of their bodies, such as standing on a chair, bending down, leaning forwards and backwards, and turning around.

Walking was mentioned earlier as a high level example of matching. In an experimental study (Slater, Usoh and Steed, 1995) we found that subjects who walked through a virtual environment using a "walking in place" technique reported a higher sense of presence than those who navigated the environment using a pointing device We speculate that this relationship was due to the greater match between optical flow and proprioception for the walking technique compared to use of a hand held pointing device for navigation.

(d) Extensiveness

Hendrix and Barfield (1996b) carried out experimental studies to examine the impact of sound on subjective presence. In one study spatialized sound was introduced or not into a visual VE. In the second study, the comparison was between non-spatialized sound and spatialized sound. In each case there was a significant effect on presence - spatialized sound led to a higher reported presence than both no sound and non-spatialized sound.

(e) Plot

We know of no study that directly attempts to examine the influence of plot in the sense of "story line". However, the study by Welch et. al. (1996) included interactivity as one of the independent variables. Again, interactivity, in the sense of whether or not the subjects drove the simulated vehicle or merely observed the VE, had a positive association with reported presence.

2.4 The Utility of Presence

Why is it important to study presence? One answer is simply to do with a strategy for research. The distinguishing feature of immersive VEs (IVEs), compared with exocentric desktop display systems, is that they afford a sense of presence. This therefore provides a direction for research - if we can find important factors that contribute to presence, then this can guide the future of the technology.

Another answer is to do with the utility of presence itself, and its relationship to "task performance". For example, this is stated, for example, by Welch et. al. as one of the reasons for studying presence (though not necessarily the main reason). Our view is that there is no reason to expect a positive association between presence and task performance. Presence is hardly the most important factor in this regard; the quality of the user interface is, for example, a crucially determining factor. In our view presence is important because the greater the degree of presence, the greater the chance that participants will behave in a VE in a manner similar to their behaviour in similar circumstances in everyday reality. Hence if a VE is being used to train fire-fighters or surgeons, then presence is crucial, since they must behave appropriately in the VE and then transfer knowledge to corresponding behaviour in the real world. There could obviously be cases where presence would diminish performance, just as being present in a situation in real life using a machine with a poor "user interface" similarly affects performance adversely.

The utility of immersive VEs in psycho-therapy relies very much on this connection between a similarity of behaviour in real and virtual environments, as has been pointed out by Strickland (1996). Responses such as acrophobia (Rothbaum et. al., 1995), claustrophobia, and fear of flying (Hodges, et. al., 1995) have been observed in immersive VEs. Clearly, these are excellent examples of behavioural presence (without presence the psychotherapy would not be possible) and yet are poor examples of "task performance", for example, the task of "travelling in an airplane", on the part of the subjects involved.

In (Slater, Linakis, Usoh, Kooper, 1996) we explored the relationship between immersion, presence and performance. This concerned a task involving comprehension and memory of a complex 3D object, events in relation to that object, and the subsequent reproduction of those events in the real world. The results suggested that increased immersion in the form of egocentric rather than exocentric viewpoint, and greater vividness in terms of richness of the portrayed environment, does indeed improve task performance (other things being equal such as relevant background knowledge and ability). The study also found that reported presence was higher for egocentric compared to exocentric immersion, but that presence itself was not associated with task performance.

2.5 Comparison with Other Proposals

The most important idea that we have presented here is the idea of external, objectively measurable characteristics that lead to a capability of placing an individual inside a computer generated environment. This is what we have called immersion, and have considered immersion ideally requiring inclusive, extensive, surrounding, and vivid display systems, where there is real-time matching between proprioception and sensory data. The VE should portray a story line, in which the individual can participate and modify. On the other hand, presence is the potential psychological and behavioural response to immersion. A highly present individual should identify with the virtual body portrayed in the VE, and therefore consider him or her self as being located in the environment in which that body is portrayed. Such a highly present individual would be observed to behave in a VE in a manner similar to how they would behave in a similar environment in everyday reality.

These ideas are only a particular distillation of the approaches of others mentioned previously (Heeter, 1992; Held and Durlach, 1992; Loomis, 1992; Sheridan, 1992; Steur, 1992; Barfield and Weghorst, 1993; Barfield et. al., 1995). In particular, Sheridan (1992; 1996) proposed three orthogonal attributes that could form a scale for presence: (a) the fidelity of the multimodal displays, (b) the ability to modify sensor position, (c) and the ability to change the configuration of the environment. In the scheme proposed in this paper, (a) is an elaboration of vividness, (b) is included in the concept of "matching", and (c) in the concept of "plot". The attributes of inclusive, extensive and surrounding can be considered as additional orthogonal attributes that may be added to Sheridan's scheme.

In his response to Sheridan, Ellis (1996) points out that a required characteristic in any proposed equation purporting to describe presence, it must be possible to demonstrate iso-presence equivalence classes, where a group of factors vary in a compensatory way so as to demarcate constant levels of presence across the variation in their range. The factors in the model presented here must, in future studies, be constructed in the manner suggested by Ellis, towards the achievement of a useful scale capable of leading to a valid measure of presence.

We take issue, however, with Ellis' remarks concerning the possible dis-utility of presence in task performance, since there is an association of the notion of "presence" with "realism". Two examples are given where it is clear that a realistic visual representation of information (air traffic display, and orbital trajectories in the vicinity of a space station) could lead to deficiency in task performance compared to a distorted representation. However, first, both environments are external, seen through a "window". Our notion of presence is that it is related to the environment in which the (virtual) body of the participant is acting. It is the relation to the interior of the aircraft cockpit that is relevant for presence, not the environment that can be seen through the window of the cockpit. Secondly, presence does not imply realism. Here is where the conceptual distinction between immersion and presence is useful. The question to ask is: what display characteristics (relevant to a certain application domain) maximise presence? It may be the case that a non-realistic display enhances presence, or that the characteristics that enhance presence are not the same as those that enhance a particular type of task performance. The separation between immersion and presence allows both to be investigated, and even if it turns out that they are correlated in a particular application, this may not be due to causal connection.

3. Shared Environments

3.1 The Abstract Society

In *The Open Society and its Enemies*, written more than 50 years ago, the philosopher of science Karl Popper envisaged a future society where most contact between humans was mediated electronically:

> '... an open society ... may ... lose the character of a concrete group of men, or of a system of such real groups ...We could conceive of a society in which men practically never meet face to face - in which all business is conducted by individuals in isolation who communicate by typed letters or by telegrams, and who go about in closed motor cars.... Now the interesting point is that our modern society resembles in many of its aspects such a completely abstract society.'
> (K. Popper, The Open Society and its Enemies, Vol. I, 1945)

This "abstract society" foreseen by Popper in 1945 is really happening now, under the popular embracing name of "cyberspace" - a huge growth in use of the Internet, and several systems that support distributed virtual environments (for example, Carlsson and Hagsand, 1993; Macedonia, 1994; Greenhalgh, 1995).

Popper envisaged the electronically mediated society as the antithesis of collectivism (the Open Society was written as a philosophical polemic against Plato - Volume 1 - and Hegel and Marx - Volume 2). We find though that there is a contradictory trend in the development of this media. At one it increases the possibility for totalitarianism and at the same time increases the chance for personal empowerment and creativity. The electronically mediated society is more likely to be like the anarchic Cyberspace of William Gibson's Neuromancer[1] than Popper's vision. In Chapters 10 and 11, we introduce the notion of 'co-presence' or 'being together' - where several people share and meet in the same environment. First we examine how the sense of 'touch' in particular force-

---

1. William Gibson, Neuromancer, Grafton.

feedback influences togetherness, and subsequently we compare behaviour of small groups when they meet in real and virtual environments, to carry out the same task.

4. Methodology

The experiments described in the subsequent chapters of this book all have a similar form. The technique of case-control experimental studies has been used to isolate various contributory factors to presence in a systematic way. In a case-control study, subjects are randomly assigned to different groups. The groups all undergo an experiment carried out under identical conditions, except that there is a systematic variation of a set of independent factors across the groups. A 'factor' may be thought of as a variable with a discrete number of possible values. For example, suppose as in Chapter 2 there is a study of the influence of whether or not having a virtual body influences presence, then 'Virtual Body' would be a factor with two levels: 'no' (does not have a virtual body), and 'yes'. Of course, an experiment could be designed to examine the impact of various types of virtual body appearance and functionality - which would be a factor with several levels. An 'explanatory variable' is a possible contributory influence on the main variable of study (typically presence). It is a variable which is not controlled in the experimental design, but, for each subject, takes the value that it happens to take. For example, suppose that gender were not used in the experimental design (i.e., the design did not require a certain number of males and a certain number of females to be distributed in the experimental groups in a specific way). Then the distribution of the subjects by gender just occurred by chance in the sampling procss. Gender can still be used to attempt to explain variation in the response variable (presence), but is treated in a different way statistically in the analysis.

In almost all the experiments some measure of "presence" is the response (or dependent) variable, and the purpose is to see if there are differences between the measured response for the different factor levels, and also how the response varies with the explanatory variables. If there are such differences, and they are statistically significant, then - other things being equal - it can be inferred that the independent and explanatory variables account for the variation in the response - and tentatively may be regarded as a causative factor. Caution should be exercised here though - a statical association between variables, no matter how 'significant' cannot be regarded as 'proof' of cause, only as something to note and to be explained. This is the role of theory - collections of 'facts' by themselves do not constitute an explanation, but facts embedded in an explicit theory come closer to what we might mean when we say that we 'understand' something. However, the theory, if it is worthy of the name, should lead to the design of further experiments, further collection of data, with results predicted by the theory. The more consistently accurate these predictions, the greater the support for the theory.

The type of statistical analysis used for the experiments described in this book is called 'generalised linear models'. This type of analysis attempts to explain the variation in a response variable, according to variation in a number of factors and independent variables.

Suppose an experiment consists of two factors, A and B. Factor A has m levels. The experimental design is such that p individuals (subjects in the experiment) are assigned to each combination of levels of A and B. Therefore the experiment requires $m*n*p$ subjects, organised as an $m*n$ table, with p subjects in each cell.

In this experiment the response variable is y, and yijk is the value of the response, of the kth person in cell (i,j) (i.e., level i of factor A, and level j of factor B). Suppose there is an independent variable x (there may be many such variables). The value of x corresponding to the kth subject in the (i,j)th cell is xijk.

The values of the independent variables are treated as 'fixed' (they are not random variables), and the response variable is considered as being affected by the influence of all the explanatory variables, plus a random error, which has some specific probability distribution, such as the Normal (or Gaussian) distribution with zero mean.

A 'linear model' is such that the expected (or average) value of yijk is a function that is linear in the (unknown) coefficients that express the relationship between y and the explanatory variables. In other words:

$$E(y_{ijk}) \ = \ \mu + \alpha_i + \beta_j + \gamma_{ij} + \phi_{ij}x_{ijk}$$

where the coefficients are considered as unknown constants, and of course the x-values are known. The coefficient α_i is 'the effect of being in the ith level of factor A'. Similarly β_j is 'the effect of being in the jth level of factor B'. γ_{ij} is the 'interaction effect' between the ith level of factor A and the jth level of factor B. This means that the impact of B on the response is different depending on which level of A is in operation (or vice versa). Note also that the regression slopes can also be different across the levels of the factors.

The hypothesis usually of interest are that the coefficients, except for the 'grand mean' μ, are all equal to zero. If not rejected, this hypothesis simply means that the independent factors and explanatory variables have no influence on the variation of y, which just 'randomly fluctuates' around the grand mean. Of course, each of the sets of coefficients can be tested separately (the impact of A, the impact of B, and the impact of x).

The coefficients can be estimated by least squares, and there is a well-worked out statistical theory based on the errors of the model (the difference between the 'true' values of y and the 'expected' or 'mean' values of y) having a Normal distribution with zero mean and constant variance for all i,j, and k. This is called a 'linear regression model' or in the particular context above 'two way analysis of covariance'.

A more complex situation occurs when the experimental situation is such that y does not have a Normal distribution. In our presence experiments we usually measure subjective presence as a 'count' - the number of times a certain type of questionnaire answer is given in a total of n questions. Here y is a count, between 0 to n, and under the hypothesis of no impact of the explanatory and independent variables on y, would have a binomial distribution. Here the relationship between y and the 'linear model' above is more complex. We let

$$\eta_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \phi_{ij}x_{ijk}$$

where η is called the 'linear predictor'.

Then the relationship between y and the linear predictor is set as:

$$E\left(\frac{y_{ijk}}{n}\right) = \frac{1}{1 + e^{-\eta_{ijk}}}$$

This is called the 'logistic function'. The advantage it has is that the predicted values of y are constrained to be in the range 0 to n, which would not be the case if normal regression analysis were used. This is a particular instance of a 'generalised linear model' - that is, where the mean of the response variable is not identically equal to the linear predictor, but mediated through some other function (called the 'inverse link function'). More will be said about this in the text of the corresponding experiments.

5. Conclusions

In the remainder of this book we will consider a number of experiments that try to obtain measures of presence, and co-presence (togetherness), and to related these measures to a number of possible explanatory factors. First we look at the relationship between presence responses and people's approach to constructing their internal world models through the representation systems that they use. In that same experiment we also considered the impact of having a 'virtual body' or avatar. Next we consider what happens when people enter virtual environments 'embedded' in virtual environments to the nth degree. Do they become more and more 'entranced' and therefore more present? In the same experiment we examine a number of other potential contributory factors, and continue to examine the influence of the representation systems. In Chapter 4, we consider a specific instance of 'vividness' - the power of shadows in a virtual environment to influence performance in a projectile aiming task, and on presence. We also discuss an 'objective' measure of presence, based on measured behaviour rather than questionnaires. In Chapter 5, we consider in more detail the issue of the relationship between immersion and performance. In Chapters 6-9 we consider a number of ways in which we have tried to understand the relationship between proprioception, natural body movement, and presence. Included is description of a method for moving

through a VE based on 'almost really walking'. In Chapters 10 and 11 we consider 'togetherness' - first an experiment on haptic feedback, and next on small group behaviour.

Computers were once remote and sacred objects to be seen only through glass and serviced by a priesthood of operators and programmers. Over the years they have become closer and closer to human beings, expanding to the mass of workplaces and homes, providing everything from accounting to entertainment. Now with immersive VEs they are beginning to supply us with new places to inhabit and share, determining our very sense data, resulting in new bodies, new powers. It is our optimistic belief that as we become more and more intertwined with computers they can become more and more liberating, the science fiction presented in novels such as Gibson's *Neuromancer* and Stephenson's *Snow Crash* [1] are becoming reality before our eyes. The role of experimentation in helping to build understanding of people's experience and behaviour within these emerging new virtual worlds is important, and it is an additional goal of this book to get others interested in following this path.

---

1. Neal Stephenson (1993) Snow Crash (Bantom Books, paperback).

Representations Systems and the Virtual Body

Summary

This chapter discusses factors that may contribute to the participant's sense of presence in immersive virtual environments. We distinguish between external factors, that is those wholly determined by the hardware and software technology employed to generate the environment, and subjective factors, that is how sensory inputs to the human participant are processed internally. The therapeutic technique known as neuro-linguistic programming (NLP) is used as a basis for measuring such internal factors. NLP uses the idea of representation systems (visual, auditory and kinesthetic) and perceptual position (egocentric or exocentric) to code subjective experience. The paper also considers one external factor, that is how the virtual environment represents a participant - either as a complete body, or just an arrow cursor that responds to hand-movements. A case-control pilot experiment is described, where the controls have self-representation as an arrow cursor, and the experimental group subjects as a simple virtual body. Measurements of subjects' preferred representation systems and perceptual positions are obtained based on counts of types of predicates and references used in essays written after the experiment. These, together with the control variable (possession/absence of a virtual body) are used as explanatory variables in a regression analysis, with reported sense of presence as the dependent variable. Although tentative and exploratory in nature, the data analysis does suggest a relationship between reported sense of presence, preferred representation system, perceptual position, and an interaction effect between these and the virtual body factor.

1. Introduction

A virtual environment (VE) is an environment created by the interaction of a human participant with a world displayed by computer. The displays provide information in the visual, auditory, and kinesthetic (including tactile and force-feedback) modalities. In *immersive* VEs (IVEs) sensory input to the human from the external world is, ideally, wholly provided by the computer generated displays. We use the term *display* here to incorporate outputs that provide consistent inputs (ideally) to all of the human senses; thus there are visual, auditory and kinesthetic display systems.

In an IVE the human participant provides an egocentric point from which the environment can be described (Ellis, 1991). This point determines a visual point of view, an auditory location, and a kinesthetic frame of reference from which the environment can be displayed by the computer. The human participant, constructs the world through perception of these displays, and operates in an extended virtual space, created by the interaction between the human perceptual system, and the computer generated displays. This affords the possibility of participants maintaining a *sense of presence* in the VE, that is the (suspension of dis-) belief that they are in a world other than where their real bodies are located. This is the unique possibility that IVEs offer: just as computers are general purpose machines, IVEs may be considered as general purpose presence-transforming machines.

Steuer distinguishes between presence and telepresence - the former relating to "being there" in the immediate physical environment, and the latter as the "*mediated* perception of an environment", that is presence in an environment constructed by a communications medium (Steuer, 1992). Steuer then defines "virtual reality" in terms of telepresence: "A virtual reality is defined as a real or simulated environment in which a perceiver experiences telepresence". In this paper we use the term presence to refer to Steuer's notion of telepresence.

Recent papers discussing presence concentrate essentially on the external properties of the IVE that are likely to contribute to presence (Held and Durlach, 1992; Loomis, 1992; Sheridan, 1992; Zelter, 1992; Heeter, 1992; Steuer, 1992). By *external* we mean features of the IVE system itself: how it is presented to the human subject, and how the subject interacts with the VE. Such external factors are wholly determined by the Virtual Environment Generator (VEG), that is the technology (hardware and software) that drives the displays and interaction devices. Such factors include, for example:

> (a) High quality, high resolution information being presented to the participant's sensory organs, in a manner that does not indicate the existence of the devices or displays. We include here Steuer's notion of vividness, "the ability of a technology to produce a sensorially rich mediated environment".
> (b) The environment that is being presented to the participant should be consistent across all displays;
> (c) The environment should be one with which the participant can interact, including objects and other actors that spontaneously react to the subject;
> (d) The self-representation of the participant, that is the participant's "virtual body", should be similar in appearance to the participant's own body, respond correctly, and be seen to correlate with the movements of the participant;
> (e) The connection between participant's actions and effects be simple enough for the participant to model over time.

In this paper we concentrate on *internal* subjective factors: how do the different ways that people have of processing information affect their responses to a VE, in particular their *reported sense of presence* in that environment? This is important in understanding the impact of the technologically bound external factors on different individuals. As Steuer (op. cit.) notes, "... virtual reality resides in an individual's consciousness; therefore, the relative contribution of each of these dimensions to creating a sense of environmental presence will vary across individuals."

We emphasise *reported* sense of presence, since people's reporting of the experience was not always the same as would be predicted by an observer noting their behaviour (Slater and Usoh, 1992). For example, a person could report that they had a relatively low sense of "being there" in the VE, yet during the experience manifest symptoms of fear, and verbally mention such fear, when faced with a fall over a precipice in the virtual world.

In Section 2 we discuss a methodology for accessing dimensions of a person's subjective experience that might be related to their sense of presence in a VE. The ideas are taken from the method known as neuro-linguistic programming (NLP) (Bandler and Grinder, 1975, 1979; Dilts et. al., 1979; Lankton, 1979; DeLozier and Grinder, 1989). In Section 3, we discuss those details of our experimental study relevant to this paper, and the methods for measuring these dimensions of subjective experience. In Section 4 we use multiple regression analysis as a data analysis technique to explore the relationships between the subjective factors and the reported sense of presence. The conclusions are then given in Section 5.

2. Subjective Experience

The unorthodox therapeutic model known as Neuro-Linguistic Programming (NLP) (see, for example, Dilts, 1979) was considered as a candidate for understanding people's responses to IVEs, since it is based on the two central ideas of *representation systems* and *perceptual position*.

2.1 Representation Systems

The NLP model claims that subjective experience is encoded in terms of three main representation systems, Visual, Auditory and Kinesthetic (VAK). The Visual system includes external images and remembered and constructed internal images. The Auditory system includes external sounds, and internal remembered and constructed sounds. It also includes internal dialogue, that is the person talking to himself on the inside. The Kinesthetic system includes tactile sensations, the sensations caused by external forces acting on the body, and also emotional responses (which are reduced to specific patterns of internal tactile and haptic sensations).

Practitioners of the method claim that people have a tendency to prefer one representation system over the others, at least in a given context. For example, a person may reason out problems by thinking mainly in images, whereas another may use mainly internal dialogue (the so-called, auditory-digital system). This is important in therapeutic practise, for if the client presents his or her problem using visual predicates (for example "I can't *see* my way ahead") whereas the therapist responds with a kinesthetic metaphor ("You have to *get a grip* on yourself"), the mismatch between representation systems may hinder effective communications between the two. There is no notion, however, that a person only thinks in one particular mode.

2.2 Perceptual Position

When a person represents a memory of some event, for example, a traumatic event, a crucial aspect of the representation is the *perceptual position* from which this is re-experienced. For example, with respect to the visual imagery, a person might internally see the event from the same perspective that they had at the time, as if they were there again, seeing it from a personal perspective. This is called an *associated* perspective, or from the *first position*. Alternatively, they might represent the event in such a way that they "see themselves" in it, either from the point of view of another person (*second position*) or from an abstract, non-personal point of view (*third posi-*

tion). Therapeutically, this can have a crucial impact on the affect attached to the event. Remembering a traumatic event from first position is quite different from remembering it from third position where the client sees him or herself as another actor in the situation.

The perceptual position is logically orthogonal to the representation system. For example, a person might say:

(a) I feel it as soon as I enter the room.

(b) You feel it as soon as you enter the room.

(c) It is felt as soon as the room is entered.

All three statements employ a kinesthetic predicate ("feel"). Statement (a) is made from first position - it expresses the information from the personal standpoint of the subject. Statement (b) is from second position - it purports to describe your reaction. Statement (c) is from third position - it describes the situation from an abstract point of view: "it can be felt" to which we can respond "it can be felt by whom specifically?" indicating an unspecified pronoun as the subject of the clause. Table 1 illustrates the orthogonality of representation system and position.

Table 1

Orthogonality of Representation System and Perceptual Position

	Visual	Auditory	Kinesthetic
First position	I see that it is OK	I say to myself that it is OK	I feel that it is OK
Second position	You see that it is OK	You say to yourself that it is OK	You feel that it is OK
Third position	It can be seen that it is OK	It can be said that it is OK	It is felt that it is OK

The theory and practise of NLP includes many more aspects than those mentioned here; we concentrate only on what is essential for our analysis. Also, the evidence for the NLP approach is largely anecdotal, practicioners implicitly accepting its correctness by virtue of their successes in clinical (and other) applications. As such it cannot be said to be part of normal science. In fact NLP has had a critical response from the academic psychology and counselling community. Some papers have found no support for the NLP theories (Monguio-Vecino and Lippman, 1987), some have found support (Buckner, et. al., 1987), one has criticised the methodology of those

who have found support (Shapley, 1987) and another has criticised the methodology of those who have not found support (Einspruch and Forman , 1985).

We are not taking a standpoint here on the validity of the NLP claims, but rather using the model as a basis for the formation of testable hypotheses of interest in our IVE research. The NLP approach seemed to be particularly interesting, since its major variables (representation systems and perceptual position) are related to precisely those factors distinguishing IVEs from other forms of human-computer interaction - the provision of a first person experience based on inputs to the three major sensory channels. The utility of the NLP method has been noted by others in the context of virtual environments (Pimentel and Teixeira, 1993).

2.3 Application to Presence in IVEs

The following hypotheses relate only to the particular VEG used in our experiment (see §3.1). They are, that, other things being equal:

(a) A person whose primary representation system is Visual, is more likely to experience a sense of presence than someone whose primary system is Auditory or Kinesthetic. The IVE experience, with the equipment we are using, is almost entirely a visual experience: most of what the VEG presents to the subject is visual information. Anything else that the person experiences, is attributable to the mental processing of that person, rather than what is actually presented to them by the VEG.

(b) A person who tends to process information more from the first position, is more likely to experience a sense of presence than those who tend to process more from the second or third position. The first position relates to "my" direct experience - the third position is essentially a meta-position relating to statements *about* the experience rather than the experience "first hand". The second position relates to how "you" would respond to the experience, rather than to how "I" am responding. Both second and third positions remove the subject a further step away from the direct experience itself.

The relationship of the Kinesthetic representation system to presence is a priori unclear. Although the VEG presents mainly visual information (with a small amount of associated sound) there is a necessary K component to the experience. Recall that K includes both tactile sensing and kinesthetic movement and forces. There is no tactile sense provided by the VEG, nor is there any force feedback in the system we are using. However, in order to operate in the environment the subjects must carry out physical movements that are similar to those necessary in real life to achieve a similar outcome - turning their heads to change their direction of gaze, bending down and moving their hand in a realistic manner to pick up objects. Also, however, the subject must carry out certain kinesthetic actions in a way quite contrary to everyday experience - such as moving forwards or backwards by pressing a button on a 3D mouse, and navigating by hand orientation. (This in itself is not entirely strange for people who use vehicles such as automobiles or cycles).

There is a further complicating factor with regard to the K representation system. One of the important goals of our experiment was to assess whether possession of a virtual body (VB) makes any difference to the sense of presence. Possession or absence of a VB is an external factor, related to point (d) of Section 1. Now those with a VB would see their virtual hand and arm move in conjunction with sensed kinesthetic movements. They could look down and see their feet on the ground, and their feet and whole VB would change orientation in response to a changed orientation of the real body. Those without a body would only see a disembodied 3D arrow responding to their hand movements. This in itself may be thought of as a very impoverished VB. What is the interrelationship (if any) between the possession/absence of a VB, and the K representation system? This was unknown before the experiment and analysis of the data (in fact the question was not posed at that time).

3. The Experimental Design

3.1 Organisation of the Experiment

Twenty graduate Computer Science students studying human-computer interaction were invited to take part in an experiment on "virtual reality". Seventeen actually took part, of whom nine had been assigned to the experimental group and the remainder to a control group. The experimental group subjects operated in the IVE with a VB, whereas the control group operated with a disembodied three dimensional arrow cursor that responded to changes in (right) hand orientation.

The experiments described in this paper were implemented on a DIVISION ProVision system, a parallel architecture for implementing virtual environments running under the dVS operating environment. The ProVision system consists of a DIVISION 3D mouse (the hand held input device), and a Virtual Research Flight Helmet™ as the head mounted display (HMD). Polhemus sensors were used for position tracking of the head and the mouse. Scene rendering is performed using an Intel i860 microprocessor (one per eye) to create an RGB RS-170 video signal which is fed to an internal NTSC video encoder and then to the displays of the Flight Helmet™. These displays (for the left and right eye) are colour LCDs with a 360×240 resolution and the HMD provides a field of view of about 75 degrees along the horizontal with a consequent loss of peripheral vision. The 3D mouse is held in a similar way to a gun. There are three thumb buttons, and a trigger. The left-most thumb button controls forward movement in the VE, the direction being determined by hand orientation. The right-most thumb button causes backwards movement - the middle thumb button was not used in the experiment. The trigger button, activated by the first finger, is used to select an object, by intersecting the object with the 3D arrow cursor, or hand, and then pulling the trigger.

Those in the experimental group saw a representation of their hand, and their thumb and first finger activation of the buttons would be reflected in movements of their corresponding virtual fingers. The hand was attached to an arm, that could be bent and twisted in response to similar movements of the real arm and wrist. The arm was connected to an entire body representation, complete with legs and left arm. If the subject turned his or her real head

around by more than 60 degrees, then the virtual body would be reoriented accordingly. So for example, if they turned their body around and then looked down at their virtual feet, their orientation would line up with their real body. However, if the subject only turned his or her head around by more than 60 degrees and looked down , the subject's real body would be seen to be out of alignment with his or her virtual body.

The experimental and control groups were roughly matched with respect to sex and whether or not their native language was English. These matching factors were chosen since presence might vary with gender, and might be based on cultural factors, signified by a country of origin. However, we were working in the dark, in the sense that we had no idea what factors that were under our control might influence the sense of presence.

Each subject was talked through the experiment by a guide who followed, as far as possible, a uniform script constructed and tested in advance of this pilot study. The subject was also watched by an observer who was taking notes. The view of the VE as seen by the subject was monitored by the experimenters on a TV display of which the subject was unaware. The entire proceedings, including the subject's view of the VE, were videotaped.

On entering the VE the subjects wer e placed in a long corridor with six doors. They were briefly given some training instructions regarding navigation and object selection while in the corridor. They was asked to enter each door in turn, leading to another room, where they would be instructed to carry out some tasks, or be presented with a certain scenario. These ranged from navigating through a room cluttered with furniture, reacting to objects flying towards their face or body area, being virtually "upside down", building a structure from blocks, to walking a plank over a virtual precipice. The experiment lasted between thirteen and twenty seven minutes across the range of subjects. There was no significant difference in the mean time spent in the VE by the experimental group (21 ± 6S.D. minutes) and control group (19 ± 5S.D. minutes). The experimental situation is described fully in (Slater and Usoh, 1992).

3.2 Data Collection

Data was collected in three ways - through observation and note taking, through the video recordings and through a questionnaire. This paper is concerned with an analysis of the responses to part of the questionnaire.

Immediately after exiting the virtual environment, the subjects were given a two part questionnaire to complete. Part A was to be completed immediately, and Part B after a further twenty four hours. In Part A the questions related to the subjects' experiences in the VE, such as their physical and mental reactions, their self-assessment of "being there" in the VE, information regarding their previous experience with "virtual reality", and extent of computer games playing. There were supplementary questions on their automobile and bicycle driving experience, their frequency of computer games playing and movie watching, and the extent to which they "identify" with characters or the situation in movies. The final question asked about their speed of adaptation to new surroundings, for example when visiting another country. This part of the questionnaire was collected before they left the building.

The second part of the questionnaire had only two questions:

> 1. During the experiment you entered a number of different rooms in the virtual environment and carried out some activities. Please write down as much as you can remember about each of the rooms and what you did and what happened in them.

> 2. Write as much as you want about your overall experiences in the virtual environments. Pay attention to your sense of being there or not, your physical sensations, your mental experiences, your thoughts about what happened - in fact about anything that occurs to you about what you experienced.

The purpose of the first question was to induce the subjects to think about their experiences again. The second was to obtain any additional information that we could relating to their sense of presence. We emphasise again, that this was a pilot experiment, so that we tried to find out as much information as possible in the time, without making too great a demand on the subjects. Of the seventeen people who took part in the experiment, fifteen returned the second part of the questionnaire, seven in the experimental group and eight in the control group.

3.3 The Dependent Variable

The dependent variable was the extent of "presence". We had several ways of assessing this: first, the reports of the subjects through their answers to a direct question on Part A of the questionnaire. Second, through observing their actual responses to situations of relative "danger" in the VE - such as whether or not they ducked when objects flew towards them, and whether they exhibited or verbally reported at the time any reactions to standing on a narrow plank over a precipice. Third, the extent to which they reacted to an outside disturbance (a noise) deliberately caused by the experimenters. Finally, their reaction to a "socially conditioned" response (being asked the time).

In this paper we consider only their subjective reporting of their degree of presence as given in their answers to a question in Part A of the questionnaire. The responses to the other indicators are presented in full in the earlier report (Slater and Usoh, 1992). However, briefly, the only other indicators that exhibited sufficient variation across the subjects for statistical analysis were whether or not they were observed to physically react when objects flew towards them, or when standing over the virtual precipice. The data suggested that this factor was related to the possession of a VB (all of those who did not respond were in the control group), but was not related to the subjective reporting of their degree of presence. This is interesting, since it is possible that their observable reactions to danger may have been automatic and unconscious, and independent of the opinion they formed about their sense of "being there".

The question, and frequency of response for each category of reported presence is given in Table 2.

Table 2

Subjective Reporting on Presence

To what extent did you experience a sense of being "really there" inside the virtual environment?	number
(1) not at all really there	1
(2) there to a small extent	2
(3) there to some extent	5
(4) a definite sense of being there	3
(5) a strong experience of being there	5
(6) totally there	1

Generally it is not statistical good practise to use the ordered categorical response (1-6) as a metric for the measurement of presence. Nevertheless, we do so in this spirit that we are carrying out an exploratory data analysis as part of the process of generating interesting hypotheses for later studies. This variable, therefore, was used as the dependent variable for the analysis.

3.4 The Explanatory Variables

The explanatory variables were measurements of the use of Visual (V), Auditory (A), Kinesthetic (K) expressions in their written answers to Part B question 2. In addition, there were measurements of the extent to which the subjects reported from first, second or third perceptual positions (P1, P2, P3).

The measurements were carried out as follows. The text in answer to question 2, Part B, was divided into sentences, and the number of V,A,K expressions in each sentence recorded. (Long sentences consisting of two or more main clauses linked by a conjunction, for example, were split into two). The perceptual standpoint of each sentence was also recorded. Most sentences would be from one of the three standpoints only, but a few were mixed in this regard. The measured variables were these counts as a proportion of the total number of sentences. For example, V was the number of visual words used over the number of sentences. K words were taken to include direct predicates such as "I felt that ...", but also any mention of actions involving the body, including movement ("When I moved my arm..."). No distinction was made between straightforward auditory expressions, such as reference to a sound, or auditory digital ("I said to myself ..."). Some examples are given in Figure 1.

Identification of the perceptual position of the sentence was quite simple. If the sentence referred to "my" experience, or mentioned "I", then it was classified as first position. If it referred to "you", then second position. If it had no personal pronoun then it would be third position. Third position could always be identified by asking the question "According to whom?". For example, in Figure 1(b), "it was an exciting experience" - according to whom? The writer could have said "It was an exciting experience for me" or "I found it an exciting experience" or "You would have found it an exciting experience" - but instead chose the third position form.

It is an assumption of this research that when a person uses a predicate such as "feel", or expresses him/herself from the third perceptual position, that such choices are not accidental, but reflect the person's actual processing of information in the given context.

Care was taken not to assume a certain representation system (based on our own prejudices). For example, in Figure 1(d), the subject mentions "sensation" which is often taken to mean physical sensation (K). However, sensation is a neutral word - it could have referred to the visual or auditory dimensions of the experience.

The measurements were "blind", that is without prior knowledge of the corresponding subject's answer to the question shown in Table 2. Recall also, that the text was produced about 24 hours after the IVE experience and answering Part A of the questionnaire. Finally, note that the V,A,K, P1, P2, P3 measurements are not based on the *content* of what the subject wrote, that is, what they wrote about, but only the *form*, the representations used, and the perceptual standpoint.

4. Results of Statistical Analysis

4.1 Multiple Regression Analysis

We use multiple regression to analyse the data[1]. Let y be the dependent variable presence, and y_{ij} the jth observation in the ith group (i=1,2) where i=1 is the control group (no body), and i=2 is the experimental group. Hence we have observations $(y_{ij}, V_{ij}, A_{ij}, P1_{ij}, P2_{ij}, P3_{ij})$ i=1,2 and j=1 n_i, where $n_1 = 8$ and $n_2 = 7$. It is possible to fit a standard analysis of variance least squares model allowing for the possibility of different coefficients of the independent variables as between the experimental and control groups.

4.2 Results of the Regression Analysis

We arrived at the fitted regression equations shown in Table 3. It should be noted that these equations are not to be read as predictive models, but a means of understanding the data that resulted from this experiment.

---

1.   The analysis was carried out using the GLIM statistical modeling system.

Table 3

Regression Equations
y = fitted values for the presence scale
p1 = (P1 - 0.7)
(Coefficients are rounded to nearest integer)

Group	Model
Control	$y = 8 - 27*p1^2 + 4*V - 2*A - 5*K$
Experimental	$y = 1 - 5*p1^2 + 4*V - 2*A + 4*K$

The model has a surprisingly good overall fit (the square of the multiple correlation being 0.99), and each coefficient is significantly different from zero at the 1% level, except for the coefficient of A which is significant at the 5% level. None of the other variables was significant. No term can be deleted from the model without significantly reducing the fit, according to F-ratio tests.

The model of Table 3 supports the hypotheses of Section 2.3 to some extent. It suggests that for this group of people, and under these experimental conditions, and other things being equal:

(a) That independently of whether or not the subject has a virtual body, the higher the proportion of visual predicates and references used, the greater the sense of presence, and the higher the proportion of auditory predicates and references the lower the sense of presence.

(b) For those with a virtual body, the higher the proportion of kinesthetic references and predicates the higher the sense of presence. For those without a virtual body, the higher the sense of kinesthetic terms the lower the sense of presence.

(c) The level of presence increases with first perceptual position (P1) up to a value of 0.7 and then decreases. The value of 0.7 was estimated from the data itself. The quadratic term was suggested by a scatter plot of perceptual position against the presence scale. The value of 0.7 is very close to the average level of P1 for the group as a whole (0.72). There is a different effect for those without and with the VB. For those without the VB, P1 is much more important than for those with - there is very little change in the presence score for those with the VB as P1 varies over its observed range of values.

The interpretation of the equations in Table 3 may be helped by a graphical representation. Figures 2-4 show a schematic representation of the presence score against the various modalities, over their observed ranges. Figure 2 shows that the V and A representation systems affect the presence scale in opposite directions, and that they are

independent of the VB. Figure 3 shows that the K representation system interacts with the VB. Figure 4 shows the interaction effect between the VB and first perceptual position.

The quadratic term in (c) might simply be an artifact of the "measurement" of presence - that the underlying "scale" of presence (if there is such) might not be a linear map of our ordinal scale. On the other hand, it could indicate that for those without the VB, that presence increases with the extent to which a person generally experiences the world from "first position", but that a person who is that way to an extreme, who never takes a dissociated standpoint, cannot allow that "suspension of disbelief" necessary in order to achieve a degree of presence in a virtual environment. This idea remains to be explored.

4.3 A Further Study

It does not follow from the regression model that the representation systems and perceptual position "determine" the degree of reported presence. If there is any relationship at all, it could be that the experience in the VE influenced the extent of the V,A,K and P predicates and references used in the subsequent write-up.

As an attempt to explore this possibility we carried out a small subsequent experiment. Six people from another stream of the Masters course were recruited and taken through the same experiment. This time, before entering the VE, they were given a preliminary questionnaire with one question, with 10 minutes to respond:

> Write as much as you like about your experience of travelling to the College today. Pay attention to your sense of the journey, your physical sensations, your mental experiences, your thoughts about what happened - in fact about anything that occurs to you about what you experienced.

After their experience in the IVE, they were given the same questionnaires as in the first round of experiments.

We were interested in whether the representation systems and perceptual positions that could be inferred from the responses to the preliminary question could be used to predict the reported degree of presence in the subsequent experiment, using the regression model of Table 3, based on the original 15 subjects. Analysis showed that this was not the case - the model, using the assessments of V,A, K and P1 from the answers to the preliminary question did not successfully predict the reported extent of presence. We realised, however, that the question was not suitable. The question seemed to evoke a "school essay" response in people, so that they wrote in a different style compared to the write-ups in Part B of the questionnaire - perhaps trying to write an interesting essay using "good" language. Maybe a different result would have been obtained had we given the subjects an unusual task to do, and then asked them to write about their responses to this.

The regression model based on the original 15 subjects was, however, more successful in "predicting" the reported degree of presence of the additional 6 subjects from the V,A, K and P1 data resulting from an analysis of their Part B answers. Table 4 shows for each individual their reported presence, the fitted value computed

from the regression equation, and the bounds for the 95% confidence interval on the fitted value. There are two cases where the observed value does not fall between the bounds, although in one of these the fitted value (5.4) is very close to the observed value (5).

Table 4

95% Confidence Intervals
for reported presence

reported presence	fitted value	lower bound	upper bound
3	3.0	2.6	3.5
5	5.4	5.1	5.6
3	2.8	2.3	3.3
4	3.0	2.6	3.4
4	4.6	3.1	6.0
3	3.2	2.7	3.2

4.4 Cautionary Note

In spite of the high level of statistical significance for the models, we must be cautious in drawing any but the most tentative conclusions. The models are fitted on the basis of few observations in relation to the number of parameters fitted. The metric for the dependent variable is without secure foundation. The subjects were not chosen at random from some well-defined population. This was after all only a pilot study, to throw some light on factors that might influence the degree of presence of a person experiencing an IVE. Nevertheless, it is an inescapable conclusion from the analysis that: "There is something going on here!". There is almost certainly some relationship between the extent of reported presence and the external factor (whether or not there is a VB) and the internal factors defined by the representation systems and perceptual positions used in writing about the experience. The model cannot show at all the direction of any causation; it could be that those who had a certain type of experience in the IVE would then be influenced to write about it using a particular pattern of representation systems and perceptual positions, rather than the latter influencing the former.

5. Conclusions

In this paper we have examined data generated by a pilot study on presence in virtual environments. We have pointed out that presence needs to be understood in relation both to internal and external factors. External factors include all those that are provided by the VEG itself, both the hardware and the software providing the interface for the human participant. Internal factors are those dimensions of the human's subjective experience relating to how information received in the IVE is processed. Such a study of internal factors is important for our understanding of what happens when people are immersed in virtual environments, and their responses to this or that aspect of the system.

We have examined two dimensions of this subjective experience, suggested by the model advanced by practitioners of neuro-linguistic programming: representation systems and perceptual position. In addition we have analysed the importance of one particular external factor, the possession of a virtual body.

It must be pointed out that there are many aspects of the data that do not allow a clear cut conclusion as far as the import of the VB factor. As discussed in our earlier report (Slater and Usoh, op. cit.), it is the case that the allocation of individuals to the experimental and control groups could not take account of suitable matching criteria since these were unknown. For example, we found that an accidental result of the allocation to groups was that there was a strong relationship between group and speed of adaptation to new environments. Nearly all of those in the experimental group for whom we had the appropriate data turned out to be people who claimed that they were fast adapters to new environments, and this seemed to be related to their reported sense of presence. Hence the results obtained in this study regarding the VB may be spurious, the result of confounding with other important factors.

The most important result of the analysis presented in this paper is that we will be able to use it for design in future work. It may also be helpful in leading towards an operational definition of presence itself. We are currently designing a predictive study, using a questionnaire to attempt to elicit preferred representation systems and perceptual position in advance of the IVE experience.
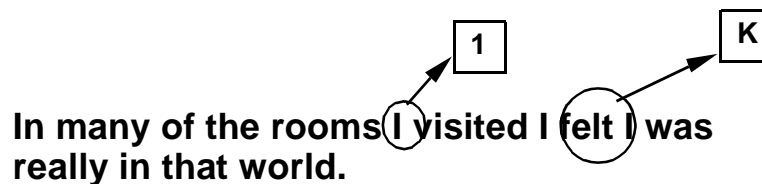
Figure 1(a)

First position, K

3

Even though the tasks and environments were quite basic, it was still an exciting experience.

Figure 1(b)

Third Position

1

A

K

I say this because initially I tried to move around the objects as if they were "real".

Figure 1(c)

First Position, A and K

**Sensation is similar to being in a dream you know is
a dream. Like you're there but you know its not real.**

2

2
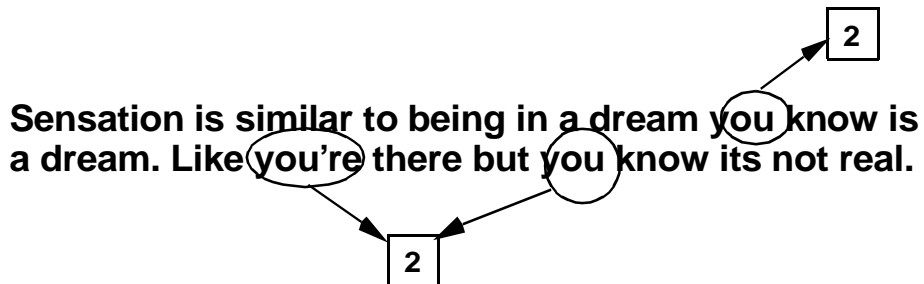
Figure 1(d)
Second Position

**The quality of the image was very grainy – so required
a lot of concentration in looking around**

3

V
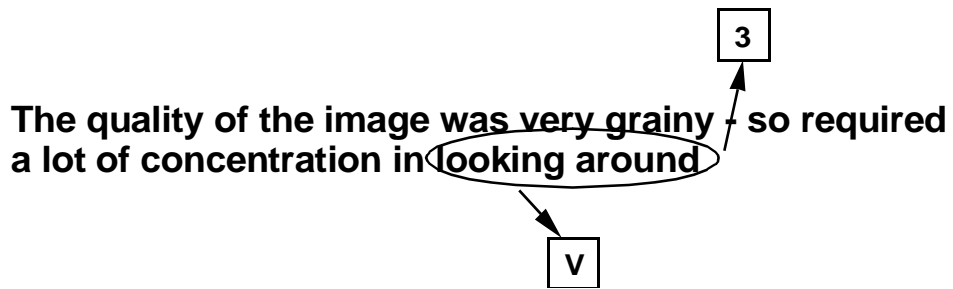
Figure 1(e)

Third Position, V
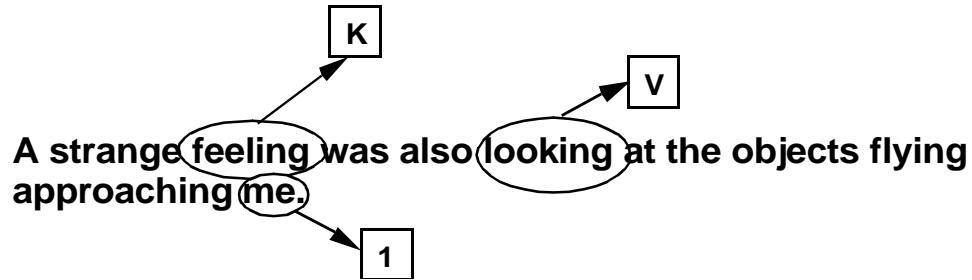
Figure 1(f)

First Position, K and V

Figure 1

Assessing Representation System and Perceptual Position - Examples
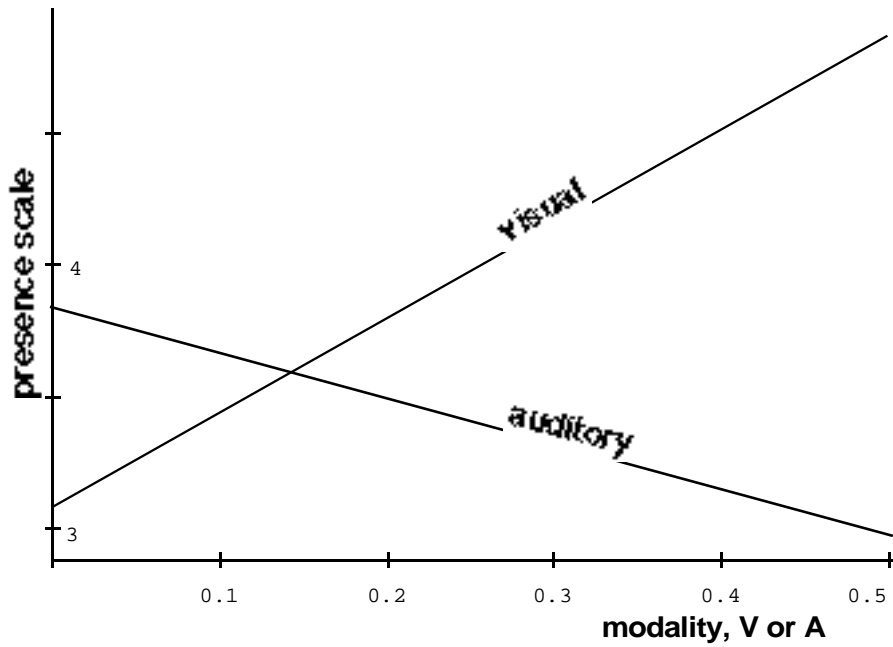
Figure 2

Presence Scale against Visual and Auditory Modalities

Using Regression Equation of Table 3

(Other variables held at their average value)
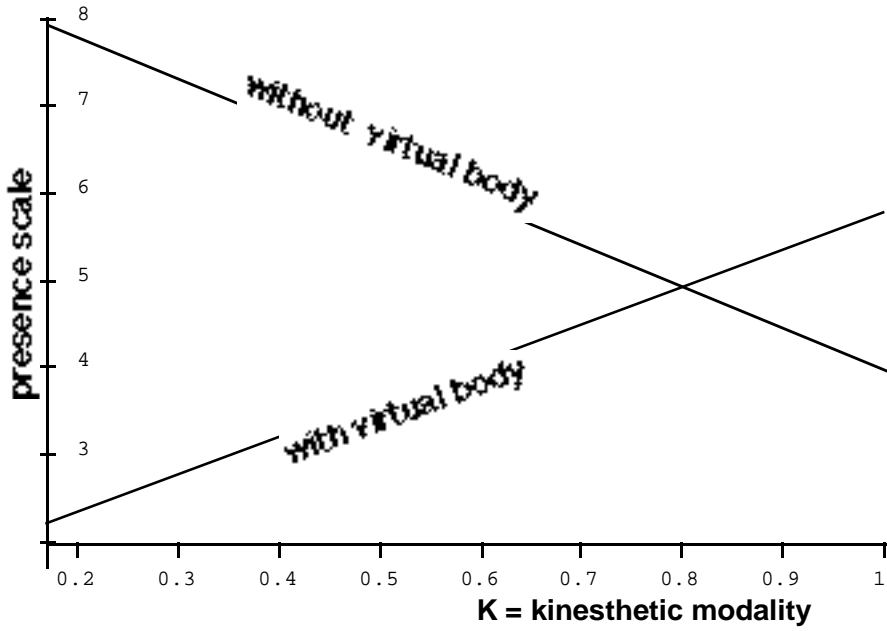
Figure 3

Presence Scale against Kinesthetic Modality

for the Experimental and Control Groups,

Using Regression Equation of Table 3

(Other variables held at their average value)
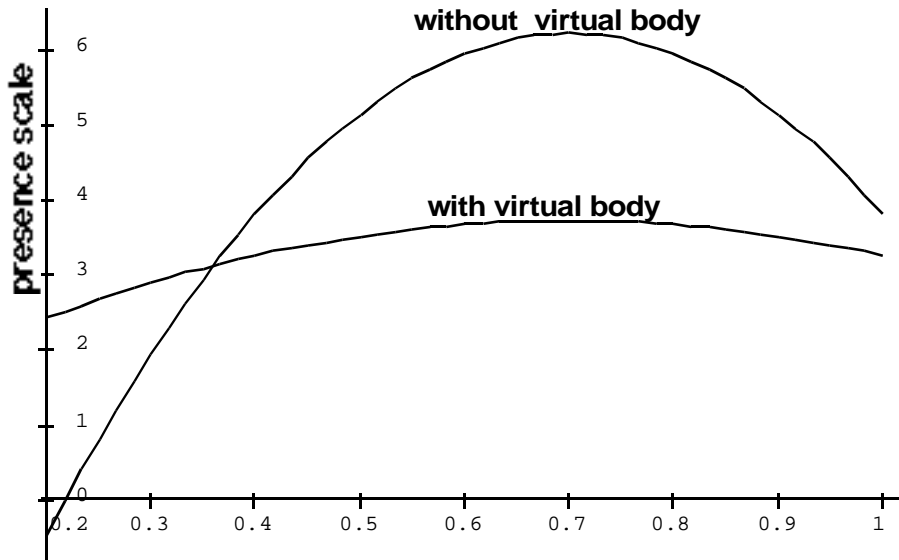
Figure 4

Presence Scale against First Perceptual Position

for the Experimental and Control Groups,

Using Regression Equation of Table 3

(Other variables held at their average value)

Depth of Presence in a Fairy Tale Setting

Summary

This chapter describes a study to assess the influence of a variety of factors on reported level of presence in immersive virtual environments. It introduces the idea of "stacking depth", that is, where a participant can simulate the process of entering the virtual environment while already in such an environment, which can be repeated to several levels of depth. An experimental study including 24 subjects was carried out. Half of the subjects were transported between environments by using virtual Head-mounted displays, and the other half by going through doors. Three other binary factors were: whether or not gravity operated, whether or not the subject experienced a virtual precipice, and whether or not the subject was followed around by a virtual actor. Visual, auditory and kinesthetic representation systems, and egocentric/exocentric perceptual positions were assessed by a pre-experiment questionnaire. Presence was assessed by the subjects as their sense of "being there", the extent to which they experienced the virtual environments as more the presenting reality than the real world in which the experiment was taking place, and the extent to which the subject experienced the virtual environments as places visited rather than images seen. A logistic regression analysis revealed that subjective reporting of presence was significantly positively associated with visual and kinesthetic representation systems, and negatively with the auditory system. This was not surprising since the virtual reality system used was primarily visual. The analysis also showed a significant and positive association with stacking level depth for those who were transported between environments by using the virtual HMD, and a negative association for those who were transported through doors. Finally, four of the subjects moved their real left arm to match

movement of the left arm of the virtual body displayed by the system. These four scored significantly higher on the kinesthetic representation system than the remainder of the subjects.

1. Introduction

This paper is concerned with the concept and measurement of presence in virtual environments (VEs) (or "virtual reality"). In a virtual environment, patterned sensory impressions are delivered to the senses of the human participant through computer generated displays (visual, auditory, tactile and kinesthetic) (Ellis, 1992). Ideally the totality of inputs to the participant's senses are continually supplied by the computer generated displays, though typically this is the case only for a subset of modalities (such as visual and auditory). In Ellis' terms, this provides the possibility of an egocentric frame of reference, where the self-representation of a participant in the VE coincides with the viewpoint from which the virtual world is experienced.

Sensors placed on the body of human participants map real body movements onto corresponding movements of their self-representation in the virtual world. We call this self-representation a virtual body (VB). We call a computer system that supports such experience an "immersive virtual environment" (IVE). It is immersive since it immerses a representation of the person's body (the VB) in the computer generated environment.

Immersion can lead to *presence*, the participant's sense of "being there" in the virtual environment. The psychological sense of presence may be considered as an emergent property of an IVE. It is important to understand the factors that contribute to this, and a means of quantifying the concept of presence itself. It has been argued (Steuer, 1992) that the very definition of virtual reality involves presence: "A virtual reality is defined as a real or simulated environment in which a perceiver experiences telepresence"[1].

In previous work (Slater and Usoh, 1993, 1994) we distinguished between external and internal factors that contribute to presence. External factors are those supplied by the IVE system itself - such as the extent of the visual field of view, auditory externalisation (outside-the-head sensations), the degree of interactivity, the behaviour of objects in the VE, and others. These factors have been discussed by (Held and Durlach, 1992; Loomis, 1992a, 1992b; Sheridan, 1992; Zelter, 1992; Heeter, 1992; Steuer, 1992; Barfield and Weghorst, 1993), and may be summarised as:

> (a) High quality, high resolution information should be presented to the participant's sensory organs, in a manner that does not indicate the existence of the devices or displays. We include here Steuer's notion of vividness, "the ability of a technology to produce a sensorially rich mediated environment".
> (b) The environment that is being presented to the participant should be consistent across all displays;
> (c) The environment should be one with which the participant can interact, including objects and autonomous actors that spontaneously react to the subject;

---

1. Steuer distinguishes between presence relating to the immediate physical environment, and telepresence relating to presence in a VE. However, we use the term "presence" throughout.

(d) The self-representation of the participant, that is the participant's "virtual body", should be similar in appearance to the participant's own body, respond correctly, and be seen to correlate with the movements of the participant;

(e) The connection between participant's actions and effects should be simple enough for the participant to model over time.

Internal factors determine the responses of different people to the same externally produced stimuli. These concern how the perceptions generated by the IVE are mediated through the mental models and representation systems that structure participants' subjective experiences. We employed the idea of a primary representation system (whether visual, auditory or kinesthetic), together with perceptual position (egocentric/exocentric) from the therapeutic model known as neuro-linguistic programming (NLP) (Bandler and Grinder, 1979) in order to construct an empirical model that relates the sense of presence to these various factors.

This paper introduces the concept of *stacking environments*. Suppose that while in a VE participants encounter and don a virtual HMD, thus simulating the activity that accomplished their original transition from everyday reality to the VE. This would place them in a deeper level environment, in which once again they could repeat the simulated transition process. To what extent is their sense of presence correlated with the depth of environment visited?

2. Indicators of Presence

(Barfield and Weghorst, 1993) provide a conceptual framework for presence referring to factors similar to (a)-(e) of Section 1, and also potential indicators of presence that may be useful for measurement. Their indicators include subjective assessment, physiometric indicators, and virtual compared to natural world task performance (an idea developed by Loomis, 1992b). In this paper we consider mainly subjective self-assessment of presence, and use three indicators first used in an earlier work to assess the influence of a navigation metaphor on presence (Slater, Steed and Usoh, 1993). These are:

(1) The subject's sense of "being there" - a direct attempt to record the overall psychological state with respect to an environment;

(2) The extent to which, while immersed in the VE, it becomes more "real or present" than everyday reality;

(3) The "locality", that is the extent to which the VE is thought of as a "place" that was visited rather than just as a set of images.

This last is similar to the idea of Barfield and Weghorst who write that "... presence in a virtual environment necessitates a belief that the participant no longer inhabits the physical space but now occupies the computer generated virtual environment as a 'place'" (op. cit., p702).

In the neuro-linguistic programming model referred to above, there is a distinction between subjective experience that is visually, auditorially, or kinesthetically encoded as mental representations. The auditory modality includes internal verbal reasoning (internal dialogue) and thinking with sound, the kinesthetic modality includes tactile and proprioceptive sensations, together with emotions (although it is argued that the latter may be decomposed into the former). It can be argued that presence, a subjective phenomenon, may be experienced differently in the different modalities. In other words, it may be possible for a person to experience a high degree of visual "presence" associated with moving through a virtual environment, while simultaneously having a sense of presence in physical reality through kinesthetic and tactile information from sitting in a real chair. Their presence in physical reality may be further reinforced aurally through conversing with someone else. Presence as such, would therefore be a function of these three effects, the reported extent of presence mainly depending on the person's dominant mode of thinking. For example, visual dominance may lead to their reporting according to their sense of presence in the visual modality. It is this synthesized, actually reported sense of presence, that we are concerned with in this paper.

The *displayed environment* is defined as that created by the virtual reality system. Note that this environment may be one created in only a single modality, such as a purely visual environment, or a purely aural environment. Ideally, it should be created in all sensory modalities. We noted that a subject may experience presence differently in different modalities, and may therefore be simultaneously aware of a number of different environments, including internal environments (eg, day dreaming), in addition to the VE. We use the notation $p(E|Esdo3(D))$ to denote a numeric measure of the level of presence of the subject in environment E given that the displayed environment is Esdo3(D). For any individual in the displayed environment Esdo3(D), $p(Esdo3(D)|Esdo3(D))$ is the quantity of interest, and is a function of the various types of internal and external factors mentioned above.

3. Stacking Environments

An IVE system may be thought of as a general purpose presence transforming machine. This is one of the aspects of the technology that distinguishes it from other egocentrically based simulation models such as traditional flight simulators. A flight simulator offers a very high degree of presence, but only in one specific environment. An IVE offers presence in an arbitrary set of environments.

We exploit this notion of "presence transformation" in constructing the following model. We use R to denote the environment "everyday physical reality". In order to enter into a VE the subject goes through a particular procedure - such as donning a HMD, putting on a glove, and so on. We denote by T the transformation that when carried out by an individual in environment E, results in the individual being in a new environment T(E). Hence the environment entered when the subject enters an IVE directly from R is $T(R)$.

Now IVEs provide the possibility of simulation. Consider that while in $T(R)$ the subject repeats a simulation of the procedures (for example, donning a HMD) that are equivalent to T. (Of course, these cannot be identically simulated). We denote the resulting environment by $T(T(R)) \equiv T^2(R)$. This procedure can be repeated over and

over - leading to the idea of environment $T^i(R) \equiv E_i$ (say), with $E_0 = R$. We call this the process of *stacking environments*, and E_i is at *depth* i in the stack.

We can say that E_1 is "once removed" from R, E_2 is "twice removed" from R, and so on. The word "removed" is a spatial metaphor, implying distance. In what dimension is the "distance"? Clearly it is not a Euclidean spatial distance, since relative to the geometry of environment R, the subject could be positioned in exactly the same place throughout. It is our hypothesis that the dimension through which the subject moves may be regarded as corresponding to presence. More specifically, the hypothesis is that

(1) $p(R \mid E_n) < p(E_n \mid E_n)$, all n > 1,

so that

(2) $p(E_n \mid E_n)$ increases with n.

Starting from R, and applying the transformation T over and over again, forces the subject on a trajectory through the "presence" dimension. If we can establish experimental procedures that correspond to this idea, and produce measurements that do in fact correspond with the *depth* of the stacked environments, then the depth of environment may be used as an equivalence class for presence. Suppose p' is an experimental measurement of presence that obeys (2), in other words it is correlated with stacking depth, we can then use p' in experiments that *do not involve stacking*. For example, if we find for some environment E that $p'(E|E) = p(E_i|E_i)$, then we can say that environment E is, with respect to its "presence" inducing attributes, equivalent to an environment that is stacked i levels deep, with reference to baseline sequence of environments $E_1, E_2, ..., E_n$. This could be used as a way of classifying virtual reality systems with respect to their presence inducing qualities.

4. Experimental Design and Procedures

4.1 Factors Expected to Influence Presence

In the design of the experiment to test these ideas we considered several other factors, noted in Section 1, thought to contribute to the sense of presence in a virtual environment.

(a) Laws of Physics

In the earlier pilot experiment (Slater and Usoh, 1992), one of the factors reported by subjects as decreasing the sense of presence was the problem that the virtual world did not behave as expected - the laws of physics were violated. (This would only be relevant in environments, such as for architectural walkthrough, where expectations of conformance to reality would be high). For example, gravity did not work correctly, objects could be penetrated, and so on.

(b) Visual Cliff

The earlier experiment also suggested strongly that presence was increased when subjects were presented to a scenario where they were standing on a narrow ledge over a precipice, reminiscent of E.J. Gibson's visual cliff experiment (Gibson and Walk, 1960). Subjects exhibited strong physical reactions (legs shaking, exclamations), and later reported this aspect of their experience as increasing their sense of presence.

(c) Virtual Actors Responding to the Subject

We also found the looming effect, where subjects "ducked" in response to an object flying towards their face, influenced presence. However, this may also related to (c) of Section 1, the idea of Heeter (Heeter, 1992) that presence is increased when there are actors in the environment that seem to spontaneously react to the subject.

(d) Subjective Factors

The results of the earlier pilot experiment strongly suggested a relationship between presence and the dominant representation system of the subject, elicited from a content analysis of essays written by them after their experience in the IVE. An exploratory statistical analysis indicated, with respect to the (mainly visual) experience of the displayed environment, that the subjective rating of presence was positively correlated with increasing visual dominance, and negatively correlated with increasing auditory dominance. For those subjects who had a virtual body as self-representation, a high score on the kinesthetic modality was positively correlated with reported presence, whereas for those who had a self-representation as a 3D arrow cursor, high kinesthetic score was negatively correlated with presence. The analysis also suggested a relationship between presence and the extent to which reporting was from an egocentric or exocentric perceptual position (Slater and Usoh, 1994). However, since this analysis was post-hoc, that is based on essays written by the subjects after their experience, and given the relatively poor means of eliciting and measuring presence, these could be only be treated as tentative results.

4.2 Factorial Design

Twenty four subjects were selected by the experimenters asking people throughout the QMW campus (in canteens, bars, laboratories, offices) whether they wished to take part in a study of "virtual reality". People with many different types of employment agreed, office secretaries, maintenance workers, researchers and students. The twenty four people were randomly assigned to the cells of Table 1, which have the following interpretations:

Table 1

Experimental Design

(a) Stacking Environments Using HMD

	No Gravity	**Gravity**
Virtual Actor that is still	Visual Cliff	No Visual Cliff
	Number of levels: 2,4,6	Number of levels: 2,4,6
Virtual Actor that follows subject	No Visual Cliff	Visual Cliff
	Number of levels: 2,4,6	Number of levels: 2,4,6

(b) Going through Environments via Doors

	No Gravity	**Gravity**
Virtual Actor that is still	Visual Cliff	No Visual Cliff
	Number of levels: 2,4,6	Number of levels: 2,4,6
Virtual Actor that follows subject	No Visual Cliff	Visual Cliff
	Number of levels: 2,4,6	Number of levels: 2,4,6

Stacking Environments

The entire scenario consisted of up to six different environments. Subjects could pass from environment to environment according to only one of two methods, depending on whether they had been assigned to Table 1(a) or Table 1(b). Those (12) assigned to Table 1(a) would see a representation of a HMD while in a VE, and would learn to transport from one level to the next by donning the virtual HMD. Those assigned to 1(b) would simply go through a door to the next environment.

Gravity

Those (12) assigned to the "Gravity" column of Table 1 would discover that objects were influenced by a simple gravity model. Those in the "No Gravity" column would discover that if they picked up an object, and then let go, it would hang suspended in space.

Virtual Actor

Those (12) assigned to the "Virtual Actor that follows subject" row would find at a certain moment in the experience that there was a humanoid-like virtual actor that would be following them around wherever they moved. Those assigned to the "Virtual Actor that is still" row would discover such a Virtual Actor, but it would remain still, unresponsive to the subject.

Visual Cliff

Those in the "Visual Cliff" cells would be forced at one moment in the experience to stand on a narrow ledge over a precipice. Those in the remaining cells would not have this experience.

Stacking Level

Each subject experienced either 2, 4 or 6 different environments.

4.3 Story Line

From the experience with our earlier pilot studies there was a strong requirement to minimise contact between experimenters and subjects once the experiment had started. Subjects in the pilot study had reported as a factor that decreased the sense of presence the verbal interchange between experimenters and subjects caused by the latter giving instructions to the former. We therefore wished to construct a story line that explained itself, requiring no input from the experimenters after a short initial training period. A story line moreover had to be one that required subjects to pick up items and let go of them (to experience the gravity effect), to move from environment to environment, to experience (or not) the visual cliff, and to give some meaning to the appearance of another humanoid computer generated actor.

This could have all been framed as a set of relatively abstract tasks with no intrinsic meaning. However, an alternative idea was to mobilise childhood fairy tale memories, to weave an overall plot, that even if senseless from a conscious rational point of view, would somehow "make sense" at an unconscious level. The scenario was therefore presented as a mixture between "Excalibur" and "Beauty and the Beast". A set of swords embedded in stone were hidden in each environment. The subject was required to find the swords, pull out the one sword amongst each set that could be displaced, find a nearby "well" and put the sword down the well. On finding the correct sword, they would awaken "the Beast". If not, they would continue their search. This scenario required a high degree of interaction with the objects in the environment: in the search for the swords the subject would have to look behind objects, open cupboards, and so on. Those who were meant to experience the visual cliff, would dis-

cover a set of swords embedded in a ledge over a precipice. In bending down to extract a sword they would be forced to become aware of the precipice. The appearance of the Beast was a natural part of the story. (A short description of the various scenes is given in Appendix A)

4.4 Procedures

(a) When initially agreeing to take part in the experiment, the subject was given a questionnaire to complete, a time was booked for several days later, and they were asked to bring the completed questionnaire to the session. The questionnaire related to the neuro-linguistic programming representation systems and perceptual position. One subject did not complete this questionnaire, so that the part of the analysis requiring use of the representation systems variables, was based on 23 subjects.

(b) On arrival for the experimental session, the subject was given a sheet to read, briefly outlining the various procedures, including a warning that some people experience a degree of nausea caused by virtual reality. The subject was given the opportunity to withdraw at that point (though none did so).

(c) The experimental guide introduced the subject briefly to the HMD. The subject donned the HMD for a short training period. This offered training in picking up objects, and moving through the environment. The training lasted for a maximum of five minutes.

(d) The HMD was removed and the subject given the story line to read, and any questions were answered at that point.

(e) The HMD was donned once again, and the main experiment completed.

(f) The HMD was removed, and the subject taken to another room to complete the post-experiment questionnaire.

4.5 Equipment and Representation

The experiment was implemented on a DIVISION ProVision200 system, under the dVS operating environment. The ProVision system consists of a DIVISION 3D mouse (the hand held input device), and a Virtual Research Flight Helmet™ as the head mounted display (HMD). Polhemus sensors were used for position tracking of the head and the 3D mouse, with a sampling rate of about 30 Hz. Scene rendering is performed using an Intel i860 microprocessor (one per eye) to create an RGB RS-170 video signal which is fed to an internal NTSC video encoder and then to the displays of the Flight Helmet™. These displays (for the left and right eye) are colour LCDs with a 360×240 resolution and the HMD provides a field of view of about 75 degrees along the horizontal with a consequent loss of peripheral vision. During the experiment the frame update rate varied between 7 and 16 Hz.

The 3D mouse is held in a similar way to a gun. There are three thumb buttons, and a trigger. However, only two buttons are used. The first (trigger) finger, is used to select an object, by intersecting the object with virtual hand and then pulling in and holding the button. The object is released by releasing the button.

Navigation is achieved by pressing on the middle thumb button (in the place of the hammer in the gun analogy). Direction of movement is determined by the direction in which the hand is pointed. For example, to move backwards, the hand can point behind, and the middle thumb button depressed. Velocity is constant. A single small step can be made by a single click of the thumb button. Subjects were standing throughout the experiment, and were free to walk around within the range determined by the trackers.

All subjects saw a VB as self representation. For example, the subjects would see a representation of their right hand, and their thumb and first finger activation of the buttons would be reflected in movements of their corresponding virtual fingers. The hand was attached to an arm, that could be bent and twisted in response to similar movements of the real arm and wrist. The arm was connected to an entire body representation, complete with legs and left arm. If the subject turned his or her real head around by more than 60 degrees, then the virtual body would be reoriented accordingly. So for example, if they turned their body around and then looked down at their virtual feet, their orientation would line up with their real body. However, if the subject only turned his or her head around by more than 60 degrees and looked down, the subject's real body would be out of alignment with his or her virtual body.

4.6 Reinforcement of Environment Transitions

Those subjects assigned to Table 1(a) of the factorial design, that is who made the transition from environment to environment using a virtual HMD, encountered a representation of a HMD within the first training environment. After they had identified this as a HMD they were invited to pick it up and put it on. Upon virtually placing it on their head, the current scene scene would be replaced by a black display, and after a short delay, a new scene would be shown. This simulated an aspect of what happened when subjects put on the real HMD: at first there was a black display, and then the experimenter switched on the HMD. The virtual HMD is shown in Figure 1.

In Section 3 we referred to the fact that the transition between environments, accomplishing transformation T, could be simulated within a VE, but that this simulation could obviously not be an exact replication of the initial transition from R to $T(R)$. For example, the experience of weight while holding the real HMD could not be reproduced for holding the virtual HMD. However, we provided additional reinforcements to the process of transition, in an attempt to stimulate, for a virtual transition, the subjective experience corresponding to a real transition.

At the start of the training period, the subject was instructed to pick up the real HMD and put it on. The HMD was, at this moment not switched on. After the subject was wearing it, at the moment that it was switched on a specific sound was played, and he or she was touched lightly on the back. Thus the transition point was marked by auditory and tactile stimulii. When the subject donned the real HMD for a second time, at the start of the experiment proper, these same auditory and tactile stimulii were generated. At each point that the subject made a transition with a virtual HMD, these same sensations were generated. The idea was to associate the moment of transition between environments with the auditory and tactile events, utilising the NLP idea of "anchoring". Anchoring is similar to classical conditioning in that it associates an internal response with external stimulii, in an attempt to evoke the internal response whenever the external stimulii are fired.

5. Instruments

5.1 Representation Systems and Perceptual Position

Prior to the experiment, subjects were asked to complete a questionnaire. The purpose of this was to attempt to elicit their primary representation system (Visual, Auditory, Kinesthetic) and perceptual position. The idea of the questionnaire was to present the subject with a number of questions with multiple choice answers, where the answers were to be ranked in accordance of appropriateness. There were typically three answers to each question, one corresponding to a V, A or K answer. For example,

You are thinking of a close friend. Rank the following in order that is most likely to correspond to your thoughts (1=most likely, 3=least likely).

The way I would think of my friend is...	Rank each answer: 1= most likely, 2 = next most likely, 3 = least likely
I mentally hear the voice and laughter of my friend.	
I feel as if my friend's presence is close to me.	
I see my friend's appearance.	

There were 11 such questions, and an individual's score on each of V, A and K would be the number of 1s (highest rank scores) for each. Note that since the subject was forced to rank the answers, one degree of freedom is lost - that is, not all of V, A and K can be simultaneously considered in an analysis.

Similarly, perceptual position was considered by questions such as:

Think of a pleasant location that you have visited, and where would like to be now. When you think about this, *how* are you thinking about it? Is it more like perceiving in your mind's eye the location as if you were seeing *a film about it*, or is it more like experiencing it *as if you were really there*?

When I think about the location ...	Rank 1= most likely, 2 = least likely
it is more like seeing a film about it	
it is more like experiencing it as if I were there	

There were six such questions, and the egocentric perceptual position was measured by the number of highest ranks (1s) given to answers corresponding to the egocentric answer.

Hence each subject, would have a V, A, and K classification as a score between 0 and 11 corresponding to the number of 1s associated with each, and a perceptual position score (P1) between 0 and 6 corresponding to the number of 1s for to the egocentric answers.

5.2 Presence

Subjective experience of presence was elicited through three questions separated from one another in the questionnaire given after the experiment:

4. Please rate *your sense of being there* in the computer generated world on the following scale from 1 to 7:

In the computer generated world I had a sense of "being there" ...	**Please tick against your answer**
1. not at all ...	1
....	...
7. very much ...	7

7. To what extent were there times during the experience when the computer generated world became the "reality" for you, and you almost forgot about the "real world" outside?

There were times during the experience when the computer generated world became more real or present for me compared to the "real world"...	**Please tick against your answer**
1. at no time	1
....	...
7. almost all of the time	7

10. When you think back about your experience, do you think of the computer generated world more as *something that you saw*, or more as *somewhere that you visited* ? Please answer on the following 1 to 7 scale.

The computer generated world seems to me to be more like	**Please tick against your answer**
1. something that I saw	1
....	...
7. somewhere that I visited	7

The presence score for a person was the total number of 6 or 7 responses from the three questions.

There were several other questions on the questionnaire, relating to the experience of nausea, adaptability in the face of new circumstances, susceptibility to vertigo, travel sickness, job, prior experience of "virtual reality", and degree of computer usage. These are not considered in this paper.

6. Results

6.1 Summary

The purpose of the experiment was to test hypotheses relating to the subjective factors (V,A,K, P1), the gravity, visual cliff, and virtual actor factors, and the stacking of environments variable, to the level of reported presence. The statistical analysis revealed:

(a) That the measured level of presence is positively associated with V and K, or negatively associated with A. This confirms the earlier result of the pilot study mentioned in Section 4(d).

(b) That the measured level of presence is positively associated with the depth of environment, when the transitions are made using the virtual HMD, but negatively associated with depth when the transitions are through doors.

(c) No other factors were statistically significant.

In the experiment, we did not attempt to control for the amount of time that the subjects spent in the experience. However, time is not at all significant in the analysis. For example, if time is substituted for depth of environment then (b) no longer holds. There was no model that we were able to find in which time emerged as a significant factor.

6.2 Logistic Regression

Dependent Variable

The dependent variable (p) was taken as the number of 6 or 7 answers to the three questions of Section 5.2. Hence p is a count between 0 and 3.

Independent Variables

These were given by the experimental design as four binary factors, gravity, cliff, virtual actor, and whether transitions were made using the virtual helmet or doors. In addition there was the number of levels visited.

Explanatory Variables

These were the V, A, K and P1 counts as discussed in Section 5.1.

This situation may be treated by logistic regression (Cox, 1970; Baker et. al., 1986), where the dependent variable is binomially distributed, with expected value related by the logistic function to a linear predictor and where N (=23) is the number of observations and where n (=3) is the number of binomial trials per observation.

Maximum likelihood estimation is used to obtain estimates of the regression coefficients. The deviance (minus twice the log-likelihood ratio of two models) may be used as goodness of fit significance test, comparing the null model with all zero coefficients with any given model. The change in deviance for adding or deleting groups of variables may also be used to test for their significance. The (change in) deviance has an approximate Chi-squared distribution with degrees of freedom dependent on the number of parameters (added or deleted).

The only factors and variables found to be statistically significant (5 per cent level) were the V, A, and K counts, and the virtual HMD and number of levels. The results are summarised in Table 2.

<div align="center">

Table 2
Logistic Regression Equations
η = fitted values for the presence scale
(Coefficients are given to 1 d.p.)
V = Visual, A = Auditory, K = Kinesthetic, counts
L = number of levels

(a) Including Auditory

</div>

Group	Model
No virtual HMD	$\eta = 2.9 - 0.6*A - 0.7*L$
With virtual HMD	$\eta = -0.8 - 0.6*A + 0.3*L$

Overall Deviance = 21.916, d.f. = 18

Chi-squared at 5% on 18 d.f. = 28.869

Deletion of Model Term	Change in Deviance	Change in d.f.	Chi-squared at 5% level
A	8.06	1	3.841
HMD.L	5.941	2	5.991

(b) Including Visual and Kinesthetic

Group	Model
No virtual HMD	$\eta = -3.1 + 0.9*V + 0.5*K - 0.7*L$
With virtual HMD	$\eta = -7.5 + 0.9*V + 0.5*K + 0.4*L$

Deviance = 19.416, d.f. = 17

Chi-squared at 5% on 17 d.f. = 27.587

Deletion of Model Term	Change in Deviance	Change in d.f.	Chi-squared at 5% level
V	10.25	1	3.841
K	4.067	1	3.841
HMD.L	6.961	2	5.991

The two models are approximately the same with respect to their goodness of fit. For a good fit, the overall deviance should be small, so that a value of less than the tabulated value is significant. For deletion of terms, the change in deviance measures how much worse the overall fit would be without the corresponding terms. Here a significant result is indicated by a large change in deviance, greater than the corresponding tabulated value.

6.3 Discussion

In the exploratory statistical analysis of the earlier pilot experiment, presence was measured by the responses to a question similar to 4 in Section 5.2. This was not satisfactory statistically, since as we stated at the time, an ordinal scale response was treated as a measured variable in the regression analysis. In the current paper we are on safer grounds, since our variable may be treated as the number of "successes" (6 or 7 scores) in three trials (the three questions of Section 5.2). However, the binomial model assumes independence between the trials, which is not obvious in the present context. The questions were each separated by two others in the questionnaire, and to a respondent, not knowing the purposes of the study, and not aware of the concept of presence, it would be reasonable to assume that answers did not directly influence one another, so that the "trials" were independent. (The correlations based on the

full 24 sets of observations are 0.46 between questions 4 and 7, 0.62 between questions 4 and 10, and 0.50 between 7 and 10).

The measure for presence used as the dependent variable requires subjects to answer at least one of the three presence questions with a high rank (6 or 7) in order to obtain non-zero score overall. This results in a large number of zeros (Table 3), and might be thought of as a reason for the negative results for factors that were expected to be associated with presence (such as the visual cliff, for example). In order to examine this, we tried an alternative measure that counted the number of ranks that were at least 5. When this was used as the dependent variable a similar model was obtained: the difference being that the K and A variables are no longer significant at 5%, whereas the V, L and HMD factors remain significant. We tried various other models using a dependent variable based on the raw scores, including the construction of a new combined score based on a principal components analysis; but in no model could we find any other significant factor.

Table 3

Frequency Table for Dependent Variable

Count	Number of 1-7 scores >_4	Number of 1-7 scores > 5
0	7	12
1	4	5
2	6	5
3	6	1
Total	23	23

In Section 6.2 we have presented in detail two models. However, there is a third worth discussing. Instead of considering V, A and K separately, we can consider the influence of their associations: for example, is the effect of a high score on each of V and K together different from, say, V being high and K being low, etc? We used as indicators of association the products V×A, V×K, and A×K. Of these three, we found that V×K leads to another model with about the same level of fit as the previous two.

Table 4

Logistic Regression Equations

including VK = V×K

Group	Model
No virtual HMD	$\eta = 0.5 + 0.2*VK - 0.9*L$
With virtual HMD	$\eta = = -4.0 + 0.2*VK + 0.3*L$

Overall Deviance = 22.510, d.f. = 18

Chi-squared at 5% on 18 d.f. = 28.869

Deletion of Model Term	Change in Deviance	Change in d.f.	Chi-squared att 5% level
VK	7.468	1	3.841
HMD.L	9.09	2	5.991

The result, shown in Table 4, indicates that the higher the VK score the higher the presence score, and in this model too, the depth of environment increases the level of presence when the virtual HMD is used for the transition between environments.

7. Presence and VK Association

In Section 2 we discussed indicators of presence. There we were concerned mainly with the psychological, subjective sense of presence - with the kind of information that is mainly obtained through asking people. However, in a situation where someone is highly present, what would we expect to observe? We argued that presence is concerned with locality, and is the extent to which a person is able to believe that they are in one particular place rather than another. Suppose that they are so convinced - what are the consequences? Certainly, a person present in an environment should respond to events in that environment. (Note, the reverse argument does not necessarily hold: a person may respond to events in an environment in which they are not present. It is certainly possible to be influenced by events in a horror movie, without having any conviction whatsoever of actually being there).

There is one object in a VE which is particularly bound with presence - this is the VB of the participant. That there is a connection between the virtual body and sense of presence may be true both logically and empirically. The logical connection follows from locality - if a person's body is in a certain locality, and they have a degree of association with that body, it is more likely that such person will "believe" that he or she is in that locality. Also, our pilot study provided empirical evidence that such a connection does exist - that the degree of presence is enhanced by having a VB compared to just an arrow cursor that responds to hand movements.

Now if high presence in an environment implies that the person will respond to events in that environment, and if the VB is a particularly important environmental object, then it follows that events connected with the VB must be particularly important with regard to indication of presence. In order to explore this idea we manipulated the VB of experimental subjects, in order to observe whether they would respond with their real physical bodies. The right hand and arm of the VB is slaved to the movements of the real right hand and arm, through the 3D mouse held in the right hand. In this experiment, the left hand and arm was programmed to mirror the movements of the right hand. For example, when a subject picked up an object, they would see both virtual hands involved in the operation, the working of the two hands together actually looking quite natural. The experimenters observed whether or not subjects matched their real left hands with their virtual left hands. In other words, would the visual information that their (vir-

tual) left hands were moving in a certain way lead to the corresponding kinesthetic activities, and therefore to the relay of the corresponding proprioceptive information?

Four subjects exhibited this behaviour for some or all of the time during the experiment - that is, they matched the movements of their real left arm to their virtual left arms. Interestingly, these subjects report no difference in their sense of presence, as measured by the responses to the three presence questions than the remaining 19 subjects, and generally seem to be no different to the remaining subjects in most other respects. However, they do differ with respect to their K scores. This is shown in Table 5.

Table 5

Mean and Standard Deviation of K score by left arm matching

$|t| = 2.756$. t on 21 d.f. $= 2.080$ at 5%

Statistic	Did not exhibit left arm matching	Did exhibit left arm matching
Mean	2.8	4.8
Standard Deviation	1.3	1.3
Number	19	4

The difference in means is significant (although beware the small number of observations in one group) and in the direction we would expect. The group that did show the matching behaviour had a mean K score that is more than double that for the remainder. We could speculate that for this group, the kinesthetic information is important enough to make sure that it conforms with their visual information. The only way that they could do this, was to move their real left arm in such a way as to ensure consistency between the kinesthetic feedback from moving their arm and the visual sensory data.

This small part of the experiment illustrated another important point: This matching behaviour occurred for each subject *immediately* they saw their virtual left hand (or not at all). It could not be influenced by the factors such as gravity - for the matching occurs at the start of the experiment before any of the high level factors could have been noticed. Only the the fundamental external factors such as quality of image and field of view, would have been experienced at this stage. It leads us to speculate that there is a deep structure of presence that is not directly conscious, but nevertheless influences behaviour in a basic way. This is in contrast to the more surface, and conscious level of presence, which is what the subject is able to articulate in answer to the appropriate questions. In an IVE both kinds of presence are obviously important. The deep level will effect a person's task performance. The second will relate to their subjective evaluation of the experience.

8. Conclusions

In this chapter we have discussed the concept of presence in virtual environments that immerse subjects into worlds created on computer displays. In particular we have presented evidence that supports the notion that a person's dominant representation system may influence their reported sense of presence, and that presence may be associated with stacking depth. Several other factors were considered, but found not to be significant.

It is possible that the reason for finding no significant associations between presence and the virtual actor or visual cliff independent variables, may be due to these being presented as solitary incidents. After leaving the VE subjects are required to make an evaluation of the entire experience, based only on these single experiences. This is important, and subsequent experimental designs should take this criticism into account. However, the visual cliff and virtual actor events always occurred at the end of the session, so that they would be more likely to be remembered and effective in influencing the final state of the subjects.

It should also be noted that although the gravity, visual cliff and virtual actor variables did not show up as significant, this could well be a consequence of the measurement of presence used. It can be argued that subjects have a certain baseline level of presence in the VE, and the questionnaire may lead to an expectation that the experimenters are looking for answers that are beyond this baseline. For example, almost all subjects carefully avoid collisions with virtual objects even though intellectually they know that there are no real obstacles. It could be the case therefore that the majority of subjects most of the time experience a strong sense of presence but do not exhibit or even later report this, because most of the time nothing out of the ordinary is happening. If they do react to a virtual actor that is walking towards them, or show visible reactions when over the virtual precipice, this is because such dramatic events call forth a visible response. Thus it is important to distinguish between signs of presence, such as reactions to extreme events, or even the matching the real left hand with the virtual one, from the underlying experience of presence itself.

The questionnaire responses are themselves another such "sign" of presence - and must be treated cautiously. It might be very interesting to carry out such an experiment in a real environment, and administer a similar questionnaire as a comparison. Presence in a real environment has been studied by (McGreevy, 1993).

The idea of stacked environments may have a practical use not discussed above. In (Fairchild et. al., 1993) there is the idea of two layers of environment: Earth where novice participants and others carry out their tasks, and Heaven, where advanced participants are able to make changes as to how Earth operates. At first thought it might seem that the different levels may reduce the sense of presence, the idea being unreal compared to everyday reality. However, the stacked environments model suggests that provided the transitions from level to level are suitably made, the existence of more than one level can enhance rather than diminish presence. In this model there could even be multiple layers, where changes at each "higher" level ripple down to affect behaviours in the lower levels.

If the method of stacking environments is generally found to contribute to increased presence, then it could be used to improve the level of presence for individuals who would normally experience a relatively low level. For example, according to our model, a person who is dominant on the auditory scale would normally experience a low level of presence in a predominantly visual environment. Use of the stacking procedure could be beneficial in such a case.

Appendix A

Brief Description of the Scenes

Scene 0

This was the training scene consisting of an empty room with cupboard and 12 inch sized cube. There was training for how to move, pick up objects, and open cupboard door. Gravity would be on or off according to the experimental cell for the subject.

Scene 1

A typical living room scene with sofas, table, and television.

Scene 2

An abstract scene with randomly scattered cubes of different sizes and colours.

Scene 3

A typical office scene with desks, swivel chairs, computers, and a filing cabinet.

Scene 4

A kitchen scene with cupboards and cooker

Scene 5

A bar scene with bar and bar furniture.

Scene 6

A cliff scene on two levels. Level 1 consisted of a suspended floor with a plank leading from it over out over level 2 below. Level 2 could only be seen when the subject was on the edge of level 1 and appreciated more when the subject was on the plank. The swords that the subject from which the subject had to choose were at the end of the plank over the precipice. Level 2 below consisted of an everyday scene containing a sofa, table and chair. Note: for those subjects who were designated not to experience the visual cliff, everything was on one level.

Figure 1 (2 photos)

The virtual HMD is being examined by a participant.

The Power of Shadows

Summary

This chapter describes an experiment where the effect of dynamic shadows in an immersive virtual environment is measured with respect to spatial perception and presence. Eight subjects were given tasks to do in a virtual environment. Each subject carried out five experimental trials, and the extent of dynamic shadow phenomena varied between the trials. Two measurements of presence were used - a subjective one based on a questionnaire, and a more objective behavioural measure. The experiment was inconclusive with respect to the effect of shadows on depth perception. However, the experiment suggests that for visually dominant subjects, the greater the extent of shadow phenomena in the virtual environment, the greater the sense of presence.

1. Introduction

We describe an experiment to examine the effect of shadows on two different aspects of the experience of immersion in a virtual environment (VE): depth perception and presence. It is well-known that shadows can significantly enhance depth perception in everyday reality (Cavanagh, Leclerc, Yvan, 1989; Gregory,1990; Puerta, 1989). Shadows provide alternative views of objects, and provide direct information about their spatial relationships with surrounding surfaces. VR systems typically do not support shadows, and yet potential applications, especially in the training sphere, will require participants to make judgements about such relationships. Even the simple task of moving to an object and picking it up can be problematic when observers cannot easily determine their own distance from the object, or its distance from surrounding objects. We introduce dynamic shadows to examine whether such task performance can be enhanced.

We have argued elsewhere (Slater, Steed, Usoh, 1994) that presence is the key to the science of immersive virtual environments (virtual reality). We distinguish, however, between immersion and presence. Immersion includes the extent to which the computer displays are extensive, surrounding, inclusive, vivid and matching. The displays are more extensive the more sensory systems that they accommodate. They are surrounding to the extent that information can arrive at the person's sense organs from any (virtual) direction. They are inclusive to the extent that all external sensory data (from physical reality) is shut out. Their vividness is a function of the variety and richness of the sensory information they can generate (Steuer, 1992). In the context of visual displays, for example, colour displays are more vivid than monochrome, and displays depicting shadows are more vivid than those that do not. Vividness is concerned with the richness, information content, resolution and quality of the displays. Finally, immersion requires that there is match between the participant's proprioceptive feedback about body movements, and the information generated on the displays. A turn of the head should result in a corresponding change to the visual display, and, for example, to the auditory displays so that sound direction is invariant to the orientation of the head. Matching requires body tracking, at least head tracking, but generally the greater the degree of body mapping, the greater the extent to which the movements of the body can be accurately reproduced.

Immersion also requires a self-representation in the VE - a Virtual Body (VB). The VB is both part of the perceived environment, and represents the being that is doing the perceiving. Perception in the VE is centred on the position in virtual space of the VB - e.g., visual perception from the viewpoint of the eyes in the head of the VB.

Immersion is an objective description of what any particular system does provide. Presence is a state of consciousness, the (psychological) sense of being in the virtual environment. Participants who are highly present should experience the VE as more the engaging reality than the surrounding world, and consider the environment specified by the displays as places visited rather than as images seen. Behaviours in the VE should be consistent with behaviours that would have occurred in everyday reality in similar circumstances.

Presence requires that the participant identify with the VB - that its movements are his/her movements, and that the VB comes to "be" the body of that person in the VE. We speculate that the additional information provided by shadows about the movements of the VB in relationship to the surfaces of the VE can enhance this degree of association, and hence the degree of presence. However, we were unable to test this in the current experiment. We do, however, consider the proposition that shadows, increasing the degree of vividness of the visual displays, will enhance the sense of presence.

2. Experiment

2.1 Scenario

The experimental scenario consisted of a virtual room, the elevation of which is shown in Figure 1. Five red spears are near a wall, but behind a small screen. Another green spear is at position G. The subject begins the experiment by moving to the red square (X), and facing the spears. The instruction is to choose the spear nearest the wall, observing from position X. Having chosen that spear, the subject moves towards it, picks it up and

returns to X. There the subject turns to the left, facing a target on the far wall. The subject must orient the spear to point approximately towards the target, fire and guide it towards the target by hand movements. The instructions were that the spear must be shot at the target, and that it must be stopped the instant that its point hit the target. Finally, the subject must bring the green spear to position X. This was repeated six times for each subject.

Prior to the start of the experiment each subject was given a sheet explaining these procedures, and the first run was for practice, the experimenter talking the subject through the entire scenario. Runs 1 through 5 were carried out by the subject without intervention by the experimenter. Between each run the subject was advised to relax with closed eyes, either with or without the head-mounted display (HMD, see below), although all but one continued to wear it during the two minutes that it took to load the program for the subsequent run. Each of the five runs were the same apart from the distances of the red spears from the wall. Also, some runs displayed dynamic shadows of the spears and the small screen, while others did not.

Eight subjects were selected by the experimenters asking people throughout the QMW campus (in canteens, bars, laboratories, offices) whether they wished to take part in a study of "virtual reality". People from our own Department were not included.

Table 1
Runs of the Experiment for Each Subject
1,2,3,4 denotes the four point-light positions of Figure 1

0 denotes no shadows

Subject	No. shadow scenes / 5	Run 1	Run 2	Run 3	Run 4	Run 5
1	1	0	0	2	0	0
2	1	0	0	2	0	0
3	2	0	0	2	3	0
4	2	0	0	2	3	0
5	3	0	1	2	3	0
6	3	0	1	2	3	0
7	4	0	1	2	3	4
8	4	0	1	2	3	4

Subject	No. shadow scenes / 5	Run 1	Run 2	Run 3	Run 4	Run 5
1	1	0	0	2	0	0
2	1	0	0	2	0	0
3	2	0	0	2	3	0
4	2	0	0	2	3	0
5	3	0	1	2	3	0
6	3	0	1	2	3	0

7	4	0	1	2	3	4
8	4	0	1	2	3	4

The design is shown in Table 1, which indicates the positions of the point-light source for those runs that included shadows. Note that of the 40 runs, 20 included shadows.

2.2 Spatial Variables and Hypotheses

The variables measured in order to assess the effects of shadows on spatial judgement were as follows:

Spear Selected. S: the spear selected from observation position X. The spears ranged from 50 cm to 90 cm from the wall, positioned with 10 cm variations. The small screen in front of the spears obscured the positions where they touched the floor, for any subject standing at position X. Also, because their distances from the wall varied only slightly, their heights, as judged from position X would look the same. It was therefore very difficult to judge which spear was nearest the wall. Variable S was the rank order of the spear chosen, where 1 would be the nearest to the wall, and 5 the furthest.

The hypothesis was that subjects would be able to use the shadows of spears on the walls to aid their judgement about the closeness to the walls, so that those runs that included shadows would result in a greater number of correct spears being chosen.

Distances from Target. C: this is the distance of the point of the spear from the centre of the target at the position that it was stopped in flight by the subject.

The hypothesis was that the subjects would be able to use the shadow of the spear in flight, especially its shadow on the target wall, to help guide the spear towards the target. Therefore, the mean distance should be less for the shadow runs than for the non-shadow runs.

D: this is the distance that the point of the spear was behind or in front of the target at the position that it was stopped by the subject.

The hypothesis is as for C, except that here we would expect a greater shadow effect since the action required to stop the spear in flight (releasing a button on the hand-held 3D mouse) is simpler than that involved in guiding the spear to the bulls eye. Moreover, at the moment the spear point touched the target wall, it would also meet its shadow.

2.3 Presence Variables and Hypotheses

In previous studies we have used subjective reported levels of "presence" based on a questionnaire. In this method subjective presence was assessed in three ways: the sense of "being there" in the VE, the extent to which there were times that the virtual world seemed more the presenting reality than the real world, and the sense of visiting somewhere rather than just seeing images. In the present study these three basic determinants were elab-

orated into six questions, each measured on a 7-point scale, where lowest presence is 1, and highest is 7 (see Appendix A). The overall presence score (P) was conservatively taken as the number of high (6 or 7) ratings amongst the six questions, so that $P = 0,1,..,6$.

Although we have obtained good results with such subjective measures before, in the shadow experiment we introduced in addition a more "objective" measurement of presence. This was achieved by having one particular object (a radio) in both the real world of the laboratory in which the experiment took place and the virtual world of the room with spears.

Just before the practice run the subjects were shown a radio on the floor against a large screen in the laboratory. They were told that they would see "the radio" in the virtual world, and that occasionally it would switch itself on. Whenever they heard the sound they should point towards "the radio", and press a button on the hand-held mouse. This would act as an "infra-red" device to switch the radio off. Before they entered into the VE the radio was momentarily switched on, deliberately not tuned to any particular channel therefore causing it to play an audible but meaningless tone. Each time that the subject entered into the VE, i.e., at the start of each run they were told: "Orient yourself by looking for the red square on the floor and the radio". The radio was placed in the VE at the same position relative to the red square as the real radio was to the position of the subject just before entering the VE.

At four moments during the experiment, always while the subject was (virtually) on the red square, the real radio was moved to one of four different positions. These were 1m apart from each other, on a line coincident (in the real world) with the small screen by which the radio was located (in the virtual world). The ordering was selected randomly before the start of the experiment. The virtual radio was always in the same place. Therefore the subject would hear the sound coming from a different location compared to the visible position of the radio. The idea is that (other things being equal), a high degree of presence would lead to the subject pointing towards the virtual radio rather than the real one. Hence we tried to cause and use the conflict between virtual and real information as an assessment of presence. Those (two) subjects who did ask about the contradiction were told "Just point at where you think the radio is". Throughout, both the real radio and the virtual radio were referred to as "the radio", deliberately allowing for a confusion in the minds of the subjects.

It is important to note that we mean "presence" in a strong behavioural sense with respect to this measurement. The questionnaire attempts to elicit the subject's state of mind. The radio method though is concerned only with their behaviour. If they pointed to the virtual radio because of a need to obey the experimenter, or because it was a matter of "playing the game", then so be it. Provided that they act in accordance with the conditions of the VE, this is behavioural presence.

Let R be the angle between the subject's real pointing direction and the direction to the real radio. Let V be the angle between the subject's virtual pointing direction and the direction to the virtual radio. Small V therefore occurs when the subject points towards the virtual radio. We use $Pa = R/V$ as the measurement of the extent to which the subject tends towards the virtual radio - a small V in comparison to R would result in large Pa. Therefore larger values of Pa indicate greater tendency towards the virtual.

There were two hypotheses relating to Pa : First, that it would correlate positively with P, and second that the greater exposure of the subject to shadows, the greater the value of Pa . Of course, we would also expect that the greater the exposure to shadows, the greater the value of P.

2.4 Representation System Dominance

A clear objection to this procedure is that it could be measuring the extent of visual or auditory dominance rather than presence. Faced with conflicting information from two senses, the resulting action is likely to depend on which sensory system is "dominant". In previous work (Slater and Usoh, 1994) we have explored the relationship between dominant representation systems and the extent of subjective presence, and have always found a very strong relationship. This is based on the idea that people differ in the extent to which they require visual, auditory or kinesthetic/tactile information in order to construct their world models, and that each person may have a general tendency to prefer one type of representation (say visual) over another (say auditory). We found that in experiments where the virtual reality system presented almost exclusively visual information, the greater the degree of visual dominance the higher the sense of presence, whereas the greater degree of auditory dominance, the lower the sense of presence.

In this shadow experiment therefore we employed an updated version of the questionnaire we used in (Slater, Usoh, Steed, 1994) which is given to the subjects before attending the experimental session. This questionnaire attempts to elicit their preferences regarding visual, auditory and kinesthetic modes of thinking. It presents 10 situations, each one having three responses (one visual, one auditory, and one kinesthetic response). Subjects are asked to rank their most likely response as 1, next most likely as 2, and least likely as 3. From this a V score is constructed as the total number of V=1 scores out of 10, and similarly for A and K. Alternatively the sums of the responses may be used. These V and A variables can therefore be used to statistically factor out the possible influence of visual or auditory dominance on the radio angles.

The hypothesis with respect to V, A and K would be that V and K would be positively correlated with presence (however it is measured) whereas A would be negatively correlated, in line with our previous findings. Note that by construction, there are only 2 degrees of freedom amongst V, A and K.

3. Apparatus

3.1 Equipment

The experiments described in this paper were implemented on a DIVISION ProVision system, a parallel architecture for implementing virtual environments running under the dVS (v0.3) operating environment. The ProVision system is based on a distributed memory architecture in which a number of autonomous processing modules are dedicated to a part of the virtual environment simulation. These processing modules or Transputer Modules (TRAMs) are small self-contained parallel processing building blocks complete with their own local memory and contain at least one Inmos Transputer which may control other specialised peripheral hardware such as digital to analog converters (DAC). Several modules exist. These include:

• the module to act as the module manager.

• the DAC module for audio output.

• polygon modules for z-buffering and Gouraud shading.

• application specific modules for the user applications.

The dVS operating environment (Grimsdale, 1991) is based on distributed Client/Server principles. Each TRAM or processing cluster is controlled by an independent parallel process known as an Actor. Each provides a set of services relating to the elements of the environment which it oversees. Such elements presently consist of lights, objects, cameras, controls (i.e. input devices), and collisions between objects. Thus, an Actor provides a service such as scene rendering (visualisation actor). Another Actor may be responsible for determining when objects have collided (collision actor) and yet another for hand tracking and input device scanning. All these Actors are co-ordinated by a special Actor called the Director. Communication between the different Actors can only be made via the Director. The Director also ensures consistency in the environment by maintaining elements of the environment which are shared by the different Actors.

The ProVision system includes a DIVISION 3D mouse, and a Virtual Research Flight Helmet as the head mounted display (HMD). Polhemus sensors are used for position tracking of the head and the mouse. The displays are colour LCDs with a 360×240 resolution and the HMD provides a horizontal field of view of about 75 degrees.

All subjects saw a VB as self representation. They would see a representation of their right hand, and their thumb and first finger activation of the 3D pointer buttons would be reflected in movements of their corresponding virtual finger and thumb. An example is shown in Plates 1 and 2. The hand was attached to an arm, that could be bent and twisted in response to similar movements of the real arm and wrist. The arm was connected to an entire but simple block-like body representation, complete with legs and left arm. Forward movement was accompanied by walking motions of the virtual legs. If the subjects turned their real head around by more than 60 degrees, then the virtual body would be reoriented accordingly. So for example, if they turned their real body around and then looked down at their virtual feet, their orientation would line up with their real body. However, turning only the head around by more than 60 degrees and looking down (an infrequent occurrence), would result in the real body being out of alignment with the virtual body.

The 3D mouse is shaped something like a gun. There is a button in the position of the hammer, which is depressed by the thumb. This causes forward motion in the direction of pointing. There is a button on each side of this central thumb button, each activated by the thumb. The left one was used to fire the spears - while this button was depressed the spear would move in a direction determined by hand orientation. The spear would stop on release of this button, and could not be activated again, thus giving the subject one chance per spear. The right thumb button was used as the "infra-red" radio switch. Corresponding to the trigger is a button for the forefinger.

This is used to pick objects - squeezing this finger button while the virtual hand intersects an object results in the object attaching to the hand. Subjects were able to master these controls very quickly.

3.2 Shadow Algorithm and Frame Rates

The shadow algorithm is described in detail elsewhere (Chrysanthou and Slater, 1995). It is based on a dynamic Shadow Volume BSP tree (Chin and Feiner, 1989), constructed from polygons in arbitrary order, that is without the necessity of a separate scene BSP tree. Shadows are created as polygons in object space. Creation of new shadows and changes to shadows are communicated dynamically to the renderer via the Director.

For reasons described below, the entire scene was small, consisting of 413 triangles, of which only 52 would be likely to influence shadow creation. The frame rate achieved without shadows was 9Hz. The frame rate with shadows, 6 to 8Hz, was not very satisfactory, but due to the particular version of the dVS software architecture in use on this machine at the time of the experiment.

Without rendering the shadow algorithm runs on this machine at a frequency of between 19 and 21Hz depending on the complexity of the view at any moment. The renderer does not however run at this frequency during dynamic changes of a virtual object, due to update problems associated with the extant implementation of the dVS dynamic geometry object. Therefore, when rendering and the associated communication time is included, the frame rate is 6 to 8Hz. dVS v0.3 maintains the concept of a "dynamic geometry object". This is a vertex-face structure representing a (possibly empty) set of polygons. The actual polygons belonging to this object can be created or modified at run time. When such a change is made to a dynamic object, there is an "update" generated that sends the object to the Director for distribution to the Visualisation Actor and then onto to the renderer.

Upon any change of a virtual object the shadow algorithm recomputes the shadow scene outputting any modified shadow polygons, i.e. any polygons that have been deleted and any that have been created. This information is transmitted to the shadow generation module which will mark deleted polygons as invisible to be re-used later by new shadow polygons. The module uses a linked list structure of dynamic objects - the shadow object. Each element in the list is a dynamic object consisting of 32 shadow polygons. This linked list structure is necessary in order to break down the entire list of potential shadow polygons into smaller chunks, rather than have one dynamic geometry object for all possible shadows, since the dynamic geometry implementation can only send updates of an entire dynamic object to the Visualisation Actor. Note that a change in one single shadow polygon will result in the communication of a complete 32-polygon dynamic object. If, unfortunately, 33 shadow polygons change, then two dynamic objects consisting of 64 polygons are communicated, and so on.

There is one important implication of this for the spatial judgement component of the experiment - obviously the spear travels more slowly when there are shadows. Without shadows the mean velocity is 92 cm/sec, and with shadows 47 cm/sec. Therefore it can be argued that differences in targeting performance might result from the velocity rather than the use of shadows. However, the effect of this can be examined statistically. With regard to the influence on presence we would argue that the slower frame rate in the case of shadows would tend to have a negative effect on presence.

4. Results

4.1 Spatial Variables

Spear Selected. Shadows made no difference at all to the selection of the "correct" spear (the one closest to the wall).

Distances from Target. Consider first C the distance of the point of the spear from the centre of the target. A regression analysis was used to examine the effect of velocity, showing that velocity within each of the shadow/no-shadow groups was did not have a statistically significant effect. The mean distance without shadows is 152cm and 115cm with shadows. However, the difference between these two is not statistically significant.

Consider next D, the perpendicular distance of the point of the spear from the wall of the target. This could be positive (spear stops in front of the target) or negative, the spear stops behind). Carrying out a within-group regression analysis to examine the effect of velocity again shows that velocity is not statistically significant. The means are -39.9cm without shadows, and 3.3cm with shadows. The standard errors are 3.6 and 3.5 respectively and the difference is significant at 5%. The medians of the shadow and non-shadow D values are -3cm and -38cm respectively.

Although the within-group velocity appeared not to be statistically significant in each case, there is still some doubt about whether the inference about better performance in the case of shadows is safe. The variation of velocity within groups was not very great (the minimum and maximum velocities were 81.6 to 99.0 for the non-shadow group, and 36.0 to 60.4 for the shadow group). Subsequent experiments should attempt to produce a greater similarity in performance between the two groups.

4.2 Presence

Subjective Presence. P is the number of "high" questionnaire scores, as a count out of 6. We therefore treated P as a binomially distributed dependent variable, and used logistic regression where N (=8) is the number of observations and n (=6) is the number of binomial trials per observation.

Maximum likelihood estimation is used to obtain estimates of the coefficients. The deviance (minus twice the log-likelihood ratio of two models) may be used as a goodness of fit significance test, comparing the null model (all coefficients are zero) with any given model. The change in deviance for adding or deleting groups of variables may also be used to test for their significance. The (change in) deviance has an approximate $\chi 2$ distribution with degrees of freedom dependent on the number of parameters (added or deleted).

Table 2

Logistic Regression Equations

η = fitted values for the linear predictor of the presence scale

A = Auditory Sum, NS = number of shadows
Standard Errors shown in brackets

Model
η = 15.0 + 0.7*NS - 9.5*A
(3.7)   (0.4)

Overall Deviance =3.454, d.f. = 5

$\chi2$ at 5% on 10 d.f. = 11.070

Deletion of Model Term	Change in Deviance	Change in d.f.	$\chi2$ at 5% level
NS	4.123	1	3.841
A	9.088	1	3.841

Table 3

Normal Regression Equations

Pa = fitted values for the angular discrepancy

NS = number of shadows

Group	Model
Visually dominant	Pa = -13.6 + 10.6*NS (3.7)
Auditory dominant	Pa = 9.427 + 0.08*NS (3.7)

Multiple Correlation Coefficient, R^2 = 0.29, d.f. = 36

Table 2 shows the result of the fit with P as the dependent variable, and the number of shadow runs (NS) and the auditory sum score (A) as the explanatory variables, across the 8 subjects. These were the only statistically significant variables found, and this supports the hypothesis that subjective presence is positively related with the shadow effect. As we have found previously, given this exclusively visual VE, the greater auditory dominance, as measured by the sum of A responses to the pre-questionnaire, the less the reported subjective presence.

Angular Discrepancy. Here we take P_a as the dependent variable and carry out a Normal regression with number of shadows (NS) and the representation system scores as the explanatory variables. NS proved once again to be significant and positively related to P_a. However, the V, A and K variables were not significant. Nevertheless it seemed important to try to rule out the possibility that the result with the angular discrepancy was simply due to visual or auditory dominance. Therefore a new factor was constructed, "sensory dominance" which has the value 1 if V>A otherwise 2. Hence this directly refers to visual or auditory dominance. The result of the regression analysis including this was interesting: for those who were visually dominant, there is a significant positive relationship between P_a and NS, whereas there is no significant relationship for those who were dominant on the auditory score. This is shown in Table 3. (It so happened that 4 of the subjects were visually dominant).

5. Conclusions

There are three main issues : First, the point of this paper is not that we have an algorithm that can generate shadow umbra rapidly in dynamically changing scenes. Even in this very small scene the rendering frame rate was no where near adequate on this particular architecutre, though its performance is excellent on standard workstations running under X11 (Chrysanthou and Slater, 1995). There is clearly a lot of work to do in the location of this algorithm in the dVS system architecture, in order to obtain maximum performance by minimising communication bottlenecks.

Second, although we have considered depth and spatial perception problems in the experiment, again, this is not the major point. It is more or less obvious, from everyday reality, and from perceptual studies that shadows do indeed enhance depth perception. Moreover, our experimental design in this regard was not ideal, since we did not control a factor (velocity) that potentially has an impact on the results.

Third, the real point of the experiment was the examination of the relationship between dynamic shadows and the sense of presence. This result is not obvious, and was motivated by the idea that presence is (amongst other things) a function of immersion, and immersion involves 'vividness'. We used two independent measures - one subjective from the post-experiment questionnaire, and the other objective, as a ratio of angles of real to virtual pointing directions. Each method gave similar results, and the two measures were significantly correlated. Moreover, we found that for those people who were more visually dominant their (angular ratio) presence increased with exposure to shadows but that this did not hold for those who were dominant on the auditory scale. Increase in the subjective presence scale was also associated with an increase in shadow exposure, but with a decrease in the auditory scale. These results also support our earlier findings regarding the importance of the sensory system preferences in explaining presence.

We suspect that much stronger results on presence would have been obtained had we been able to allow the virtual body to cast shadows. However, this was not practical given the communication bottleneck problems discussed in §3.2.

If an application does not require presence, there is little point in using a virtual reality system. If a virtual reality system is used for an application, then there is little point to this unless it can be shown that a sense of presence is induced for most of the potential participants. Should the results of our shadow experiment be confirmed by later studies then it will have been shown that the great computational expense of shadow generation is worth-while for those applications where the participants are likely to be "visually dominant".
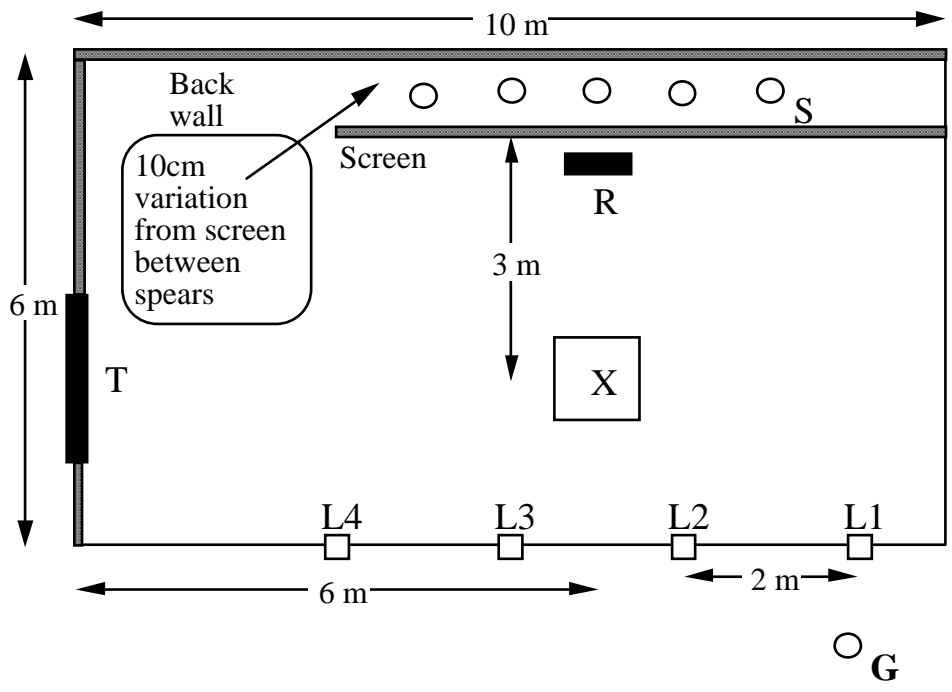
Appendix A: Presence Questions

All questions were answered on a 1 to 7 scale, not reproduced here for space reasons.

1. Please rate *your sense of being there* in the virtual reality.

2. To what extent were there times during the experience when the virtual reality became the "reality" for you, and you almost forgot about the "real world" of the laboratory in which the whole experience was really taking place?

3. When you think back about your experience, do you think of the virtual reality more as *images that you saw*, or more as *somewhere that you visited* ?

4. During the course of the experience, which was strongest on the whole, your sense of being in the virtual reality, or of being in the real world of the laboratory?

5. When you think about the virtual reality, to what extent is the way that you are thinking about this similar to the way that you are thinking about the various places that you've been today?

6. During the course of the virtual reality experience, did you often think to yourself that you were actually just standing in a laboratory wearing a helmet, or did the virtual reality overwhelm you?

Figure 1
Plan View of the Virtual Environment

10 m

Back wall

10cm variation from screen between spears

Screen
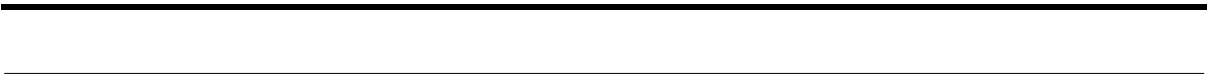
S

R

3 m

X

6 m

T

L4   L3   L2   L1

6 m

2 m

G

*
E

X - red square (judgement point)
S - spear
T - target
G - green spear

E - scene entry point
L - light (positions 1, 2, 3, 4)
R - virtual radio

Performance in a 3D Chess World

Summary

This chapter describes an experiment to assess the influence of immersion on performance in immersive virtual environments. The task involved Tri-Dimensional Chess, and required subjects to reproduce on a real chess board the state of the board learned from a sequence of moves witnessed in a virtual environment. Twenty four subjects were allocated to a factorial design consisting of two levels of immersion (exocentric screen based, and egocentric HMD based), and two kinds of environment (plain and realistic. The results suggest that egocentric subjects performed better than exocentric, and those in the more realistic environment performed better than those in the less realistic environment. Previous knowledge of chess, and amount of virtual practice were also significant, and may be considered as control variables to equalise these factors amongst the subjects. Other things being equal, males remembered the moves better than females, although female performance improved with higher spatial ability test score. The paper also attempts to clarify the relationship between immersion, presence and performance, and locates the experiment within such a theoretical framework.

1. Introduction: Is VR better than a workstation?

This chapter describes an experiment to compare task performance of twenty-four subjects with respect to immersive and non-immersive participation and interaction in a virtual environment (VE). Subjects participated in and observed a sequence of events played out in relation to a complex geometrical structure, and their task was the subsequent reproduction of those events with respect to the real world equivalent of that structure. Half the subjects participated in a visually immersive VE, and the remainder in a non-immersive VE. A secondary independent variable was the realism of the displayed environment. The task involved the reproduction of moves of

pieces in the Tri-Dimensional chess game popularised in the TV series Star Trek. One inspiration of this work was the question posed by Mizell, Jones, Jackson and Picket (1995): "Is VR better than a workstation?" For our work this question breaks down into a number of components:

(a) Cognition of Geometric Structure

We want to know whether people can benefit significantly with respect to their understanding of geometric structure when they are immersed in and interact with a VE compared to the use of conventional workstations and displays. This is important for our work on geometrical modeling, where participants are immersed in a VE in order to create free-form surfaces (Slater and Usoh, 1995). Here the issue is to gain insight as to whether immersive virtual environments (IVEs) can show any benefit with respect to their understanding of the complex geometrical structure of such surfaces and objects composed from these surfaces.

This is similar to the question posed by Mizell et. al., who aim to "experimentally assess and quantify, if possible, a difference in a user's being able to comprehend a complex three-dimensional scene between viewing it in 2-D on a workstation screen and viewing the scene via an immersive VR system...". The Mizell work showed subjects complex 3D shapes in reality, on a conventional workstation display, and through a stereo BOOM head-coupled device. The shapes were of three levels of difficulty. Each subject had to reproduce the shape in reality (using provided basic shapes) while being able to observe the baseline shapes from one of the three sources. Each subject successively used each method (reality, workstation, BOOM) in a pre-assigned randomly determined order. The response variable was time to completion of the shape. The results showed that observing the shape in reality was always the fastest method, and except for the simplest shape, looking at the workstation display was more effective than looking through the BOOM. As the authors noted, however, there were several problems with this experiment, including the fact that looking at the workstation display was a faster operation than looking through the BOOM (which could often get into awkward positions). Moreover, this experiment involved no interactive manipulation of objects in the VE.

(b) Knowledge Transfer

Here the question is whether skills or knowledge gained in a virtual environment can be successfully transferred to the real world. Suppose that a person learns to perform some task in a virtual environment, does immersion improve the chance of transfer of such knowledge to the real world? A previous attempt to study this (Slater, Alberto and Usoh, 1995) involved a number of subjects who walked through a virtual building immersively (with a head-slaved head-mounted display), and control group subjects who did the same non-immersively (looking at a monitor). The task involved finding a particular object within the virtual building. A response vari-

able was the time it took them to find this object in the corresponding real building. No significant difference was found between performance for immersed and non-immersed subjects. However, the evidence suggested that those individuals with a high sense of presence, whether in an immersive or non-immersive system, achieved better performance overall. In this study also there was no interaction with objects in the virtual environment other than walkthrough.

Wilson and Foreman (1993) considered a similar problem comparing observations from subjects in *non-immersive virtual* and the corresponding *real* environments. They concluded that "... the overall picture is one of little difference between spatial information gained from exploring the computer simulation of the building and real exploration." If this is the case for virtual non-immersive and real environments, then is not surprising that this may the case for non-immersive and virtual immersive environments. The tasks of the subjects of Wilson and Foreman, however, also did not involve interaction with the environment.

(c) Immersion and Performance Within the Virtual Environment

The studies reported above concentrated on the effects of immersion in relation to tasks later performed in the real world. It is also important to consider these effects in relation to tasks performed in a VE. Chung (1992) considered various different models of immersion on targeting of treatment beams in radiotherapy treatment planning. All subjects used a HMD but in some models the HMD was enabled for head-tracking, and in other models steering was achieved through hand-held devices. All subjects used each steering mode. The experimental study found no difference between the head-tracked steering modes and the non-head tracked modes. Pausch, Shackelford and Proffitt (1993) studied the effect of immersion on a target search task in a study where one group used a head-tracked HMD for target location, and another used a HMD with head tracking disabled with viewing controlled through the use of a hand-tracked device. The result was that the head-tracked group achieved nearly twice the speed of target location compared to the hand-trackers.

None of the studies reported above included interaction, that is, manipulation of objects in the virtual environment. Also the use of the phrase "task performance" in relation to VEs is sometimes ambiguous. There is clearly the distinction between effectiveness of task performance within the VE, and effective performance in relation to some task performed in the real world but in relation to a VE experience. In this paper we consider the latter case, and also have an experimental scenario that includes some interaction with objects in the VE.

In the next Section we provide more explicit explanations for the terms immersion, presence and the relationship between these and "performance". The details of the experiment are provided in Section 3, and results in Section 4. Section 5 provides the overall conclusions of this study and some consideration of the problems of experimental design and further work.

2. Immersion, Presence and Task Performance

2.1 Immersion and Presence

In reports of earlier studies we have made a distinction between immersion and presence (Slater, Usoh and Steed, 1995). Immersion refers to what is, in principle, a quantifiable description of a technology. It includes the extent to which the computer displays are extensive, surrounding, inclusive, vivid and matching. The displays are more *extensive* the more sensory systems that they accommodate. They are *surrounding* to the extent that information can arrive at the person's sense organs from any (virtual) direction, and the participant can turn towards that direction receiving the appropriate directional sensory signals. The notion of surrounding also includes the greater the reproduction of the natural modes of sensory presentation (visual and auditory stereopsis for example). They are *inclusive* to the extent that all external sensory data (from physical reality) is shut out. Their *vividness* is a function of the variety and richness of the sensory information they can generate (Steuer, 1992). Vividness is concerned with the richness, information content, resolution and quality of the displays. Finally,

immersion requires that there is *match* between the participant's proprioceptive feedback about body movements, and the information generated on the displays. A turn of the head should result in a corresponding change to the visual display, and, for example, to the auditory displays so that perceived sound direction is invariant to the orientation of the head. Matching requires body tracking, at least head tracking, but generally the greater the degree of body mapping, the greater the extent to which the movements of the body can be accurately reproduced.

Immersion, in our definition, also requires a self-representation in the VE - a Virtual Body (VB). The VB is both part of the perceived environment, and represents the being that is doing the perceiving. Perception in the VE is centred on the position in virtual space of the VB - e.g., visual perception from the viewpoint of the eyes in the head of the VB, an egocentric viewpoint.

Immersion, in our view, is therefore an objective description of what any particular system does provide. Presence is a state of consciousness, the (psychological) sense of being in the virtual environment, and corresponding modes of behaviour. Participants who are highly present should experience the VE as more the engaging reality than the surrounding world, and consider the environment specified by the displays as places visited rather than as images seen. Behaviours in the VE should be consistent with behaviours that would have occurred in everyday reality in similar circumstances.

Our general hypothesis is that presence is an increasing function of two orthogonal variables. The first variable is the extent of the match between the displayed sensory data and the internal representation systems and subjective world models typically employed by the participant. Although immersion is increased with the vividness of the displays, as discussed above, we must also take into account the extent to which the information displayed allows individuals to construct their own internal mental models of reality. For example, a vivid visual display system

might afford some individuals a sense of "presence", but be unsuited for others in the absence of sound (Slater, Usoh and Steed, 1994). The second variable is the extent of the match between proprioception and sensory data. The changes to the display must ideally be consistent with and match through time, without lag, changes caused by the individual's movement and locomotion - whether of individual limbs or the whole body relative to the ground.

It is important to realise that this model operates at many levels. Considering the visual display as an example, at the most basic level the important factors may be field of view, resolution, colour resolution, binocular disparity. Corresponding behaviours from the point of view of presence are those autonomic responses governed by the visual system such as vergence and accommodation (Ellis, 1991). At a higher level the realism of the content of the visual display may be considered - such as whether objects behave in accordance with physical laws. Corresponding behaviour with respect to presence may be concerned with observable gross involuntary behaviour - such as the looming effect (when an individual ducks in response to a flying object), or the experience of vertigo in response to a virtual visual cliff. At the highest level such features as the realism of the illumination may gov-

ern people's voluntary responses, such as evaluations of their sense of "being there", or the "realism" of the virtual environment.

2.2 Presence and Task Performance

It is sometimes argued that it is important to study presence because of the potential relationship between presence and performance. For example, in (Barfield, Sheridan, Zeltzer, Slater, 1995) we find:

"Not only is it necessary to develop a theory of presence for virtual environments, it is also necessary to develop a basic research program to investigate the relationship between presence and performance using virtual environments. ... we need to determine when, and under what conditions, presence can be a benefit or a detriment to performance? ... When simulation and virtual environments are employed, what is contributed by the sense of presence per se?"

The question of the relationship between presence and performance goes to the heart of why presence is important. The issue is not really that of whether presence itself enhances performance. For example, an individual's performance in word processing is usually superior using a modern point-and-click user interface than under UNIX using "vi" - not of course because of presence, but because of the former's superior user interface. In our view presence is important because the greater the degree of presence, the greater the chance that participants will behave in a VE in a manner similar to their behaviour in similar circumstances in everyday reality. Hence if an IVE is being used to train fire-fighters or surgeons, then presence is crucial, since we want them to behave appropriately in the VE and then transfer knowledge to corresponding behaviour in the real world. There could

obviously be cases where presence would diminish performance, just as being present in a situation in real life using a machine with a poor "user interface" similarly affects performance adversely.

Hence it is posing the wrong question to consider whether presence *per se* facilitates task performance. Rather presence brings into play "natural" reactions to a situation (which may or may not have something to do with efficiency of task performance) - and the greater the extent to which these natural reactions can be brought into play the greater that presence is facilitated, and so on. It isn't really a question of how good the performance is, but rather how it is grounded in presence.

We would nevertheless expect to find an association between presence and performance for some tasks - precisely those tasks that benefit from *immersion*. For the purposes of this study we postulated that increased "immersion" would lead to improved "task performance" (to be defined below in the context of this experiment). This is because the task involved comprehension and memory of a complex three dimensional structure and events relating to that structure, and we considered that performance would be enhanced by an egocentric, stereo

view based on a head-tracked HMD compared to an exocentric screen based view. Since our overall hypothesis is that both performance and presence are enhanced by immersion, we would therefore not be surprised to find an association between performance and presence.
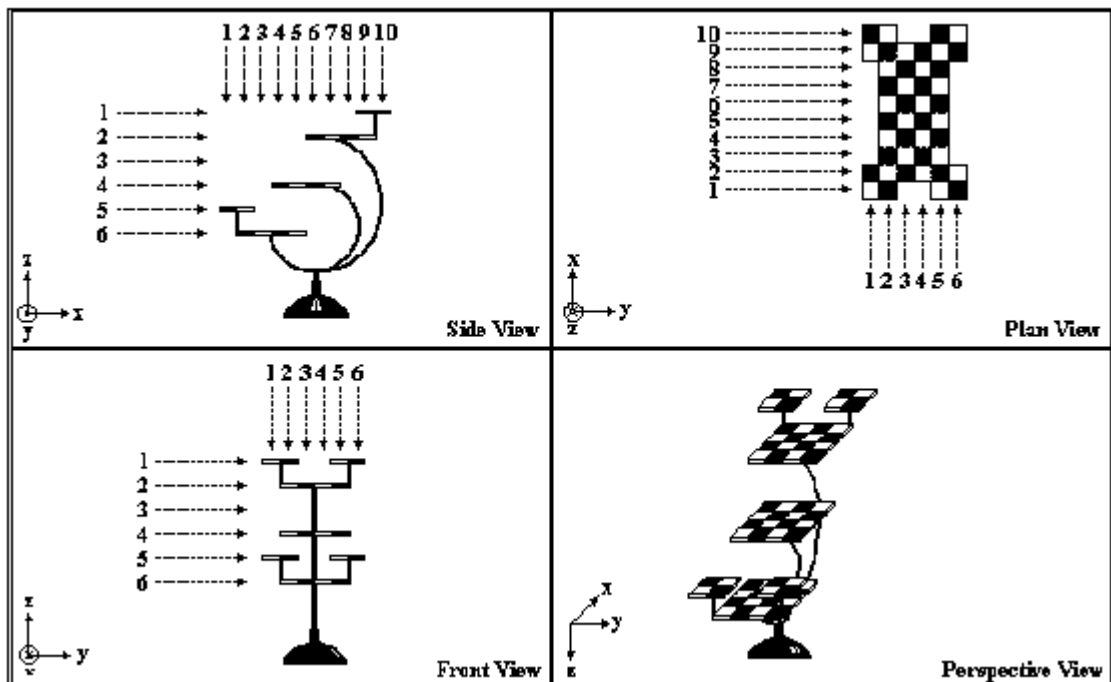
Figure 1
Tri-Dimensional Chess

3. Experiment

3.1 Background: Tri-Dimensional Chess

Tri-Dimensional Chess (TDC) is a board game which has many characteristics in common with conventional chess[1]. More specifically, it is a type of chess played on a number of boards suspended at a different heights. The pieces used in this game are the same as conventional chess and capable of the same movements, but also may be moved from one board to another. Moreover, the layout of the different boards is irregular, and the initial positions of the sixteen pieces of each side are different than in conventional chess. Finally, TDC has a set of four movable attack boards, which are also considered to be pieces, and can be moved according to certain rules (Figure 1).

TDC was chosen because it provides a complex geometrical structure and it is this complexity of the layout of the boards and the pieces, which make it a suitable vehicle for the study. The actual rules of the game were of no

_____

1.See http://www.redweb.com/chess/Variants/

importance for this experiment. Twenty four subjects were chosen for the experiment according to the factorial design of Table 1.

3.2 Factorial Design

(a) Immersion: Exocentric/Egocentric

This factor relates to the surrounding aspect of immersion discussed above. Half of the subjects were immersed with an egocentric view into a virtual environment. This was achieved using a DIVISION ProVision100, with a Virtual Research Flight Helmet and a DIVISION 3D Mouse. Polhemus Fastrak sensors were used for position tracking of the head and the mouse. The generated image has a resolution of 704x480 which is relayed to two colour LCDs each with a 360×240 resolution. The HMD provides a horizontal field of view of about 75 degrees,

and about 40 degrees vertically. Forward movement in the VE is accomplished by pressing a left thumb button on the 3D mouse, and backward movement with a right thumb button. A virtual hand was slaved to the 3D mouse - there was no virtual body representation other than this. Objects could be touched by the hand and grabbed by using the trigger finger button on the 3D mouse.

Table 1
The Number of Subjects Per Cell of the Factorial Design

Immersion:	Exocentric		Egocentric	
Environment:	7 moves	9 moves	7 moves	9 moves
Plain	3	3	3	3
Garden	3	3	3	3

The other 12 subjects experienced the VE from an exocentric view. In order to keep all conditions as similar as possible apart from egocentric or exocentric, the exocentric subjects used exactly the same system, except that they viewed the images on a TV screen. They controlled movement by the 3D mouse. This time the HMD was placed on the left shoulder of the subject so that viewpoint could be controlled with the left hand.

One condition could not be controlled - the resolution of the different displays (HMD and TV screen). The image generated from the same source as the HMD had a resolution of 704×480 which was fed to an NTSC 3.58 28 inch TV. Hence the exocentric group observed a higher resolution display.

(b) Environment: Plain/Garden

The environment factor is related to *vividness*. Half of the subjects ("garden") participated in an environment where the TDC system was located in a realistic setting. This consisted of an open field, populated by a table, a chair, a tree and small plant. The TDC board was located on the table. This model had a large horizontal plane forming the ground, and a spherical cone representing the sky. This was called the "garden" environment. All surfaces in the VE garden were appropriately texture mapped. The remaining subjects ("plain") saw the TDC game suspended in a void. Examples of these environments are shown in the colour plates.

(c) Number of Moves

Each subject had to witness the first few moves of a computer versus computer game. The number of moves was either 7 or 9, to give tasks of slightly differing degrees of complexity. The subject was responsible for initiating the sequence of moves, as well as "instructing" each consecutive move. To be more precise, the subject had to initiate this game by pressing a red button situated next to the base of the virtual TDC. As soon as the button was pressed, one of the pieces on the board would change its colour to bright red, indicating the first computer move. The move was not performed by the computer until the subject decided to "instruct" the computer to do so. To give this instruction, the subject had to touch the red piece with the virtual hand. Doing this caused the piece to leave its current position and move to a new position on the board. As soon as this piece moved to its new position, another piece on the board changed its colour to bright red. The subject had then to touch this piece in order to make it move to its new position on the board, following a predetermined path. Another piece would then in turn change to bright red, and so on. This process carried on for a certain number of moves - 7 or 9. When the subject could not find any other bright red piece on the board the sequence had finished. The subject could repeat the identical complete sequence from the beginning by again pressing the red button.

The task of the subject was to remember which pieces were moved and where they were moved to. They then had to reproduce the final state of the board on the real life TDC board from which the virtual TDC had been modeled. There was no limit to the amount of times a subject could repeat the sequence of moves. This was done so that different rates of learning between the subjects be eliminated as a source of experimental variation. The importance of feeling confident in being able to accurately reproduce the moves in real life was clearly explained to each subject, and the main experiment did not commence until the subject confirmed a high degree of confidence.

3.3 Virtual Model and Performance

The boards and pieces were modeled in AutoCAD.   There were on the average 290 vertices and 230 triangles in each chess piece. The entire board, including the board base, the bottom, middle, and top boards, the attack-boards, and the poles suspending the attack-boards consisted of 438 vertices and 344 triangles, all texture mapped. The activation button consisted of 56 vertices and 42 triangles. The garden, including a table, a chair, a plant, a tree, a ground, and a sky-dome, consisted of 2543 vertices and 1456 triangles, all texture mapped. Altogether there were a total of 7732 triangles in the garden environment and 6276 in the plain environment (consisting only of the TDC system). Further description of the process of object construction can be found in in (Linakis, 1995).

The frame rate offered on the ProVision system is not guaranteed at any particular level of performance. It varied between 15 and 20Hz depending on the complexity of the data in view at any particular time. Clearly subjects in the plain environment would have generally experienced a faster frame rate than those in the garden environ-

---

ment. This does confound the experiment to some extent since on the one hand the more realistic environment is, in the terminology of this paper, a more "immersive" one, yet its lower frame rate makes it less immersive. However, an experiment of Barfield and Hendrix (1995) found that frame rates of between 15 and 20Hz resulted in the same degree of reported "presence".

3.4 Procedures

(a) Selection of Subjects

The range of subjects was chosen to be as broad as possible in terms of their background knowledge of chess and computer literacy. The subjects varied from computer science students with previous computer and chess knowledge, to people with no previous knowledge of chess and almost no computer experience. Allocation to the cells was carried out randomly except that in cases where the subjects had previously experienced Virtual Reality they were evenly distributed in the design, so as to maintain an average of VR expertise amongst the subjects of each cell. There were 16 males and 8 females.

(b) The Pre-Questionnaire

A pre-questionnaire was given to subjects at the time of their agreement to participate. This gathered basic demographical and other information such as prior experience with chess, TDC, computers and VR.

(c) The Spatial Awareness Test

The Spatial Awareness Test (SAT) is one of four General Ability Tests which aim to measure how well a person can identify similarities and relationships in words, shapes, or numbers. The specific purpose of the SAT is to test the ability of a person to create, retain and manipulate mental images by mentally "folding" flat patterns into 3D objects (Smith and Whetton, 1988).

Each subject had to do this test, and a standard score was derived from their answers. The higher the score, the greater the ability of the subject to mentally derive 3D structure when from 2D visual input, according to the principles of these tests. The purpose of administering the test was to attempt to take into account differing background abilities in mental imagery.

(d) Introducing Subjects to the Tri-Dimensional Chess

Most of the subjects were not expected to have seen the Tri-Dimensional chess before this introductory session so that a training session should be given to each subject.

Each subject was given the same introductory talk on the TDC. They were told that the TDC has three main boards and four attack boards. It was made very clear to them that those attack boards were considered to be pieces, and that they could be part of a legal move. Finally, it was decided that the subjects should not be told that the TDC pieces are capable of the same movements as conventional chess pieces, as it would be very easy to deduce the correct position of a moved piece based on elimination of impossible moves. Subjects were therefore told that pieces can move in any one of six directions i.e. forward, backwards, left, right, up, and down (from one level to another).

Finally, the experimenter performed a number of moves on the real TDC using the pieces of one side (the Gold side), and then asked the subjects to copy the moves with the pieces of the opposite side (Silver). If a subject made an error in copying the move the experimenter would immediately report this to the subject and the corrected version of the move was performed by the experimenter and explained to the subject.

(e) The Virtual Kitchen Task

As with the TDC, the subjects were not expected to have any existing knowledge or experience in Virtual Reality. It was therefore important to familiarise them with the VR equipment before the main experimental task was to be carried out.

A virtual kitchen demonstration was chosen for this purpose since it involved an environment with which people are naturally familiar. Moreover, this environment is well designed (i.e. precise in size, and detailed in geometrical description), and is highly interactive since most of the objects can be picked up or moved.

Initially, the VR equipment was shown to the subjects. They were told how they could navigate through the virtual environment, how they could pick objects up, and release them. They were given written instructions on what they had to do in the virtual kitchen. The same instructions were repeated verbally by the experimenter. Each subject had to navigate around the kitchen, find a particular object in the kitchen, pick it up and drop it, after having taken it to another part of the environment. The particular object turned its colour to bright red when touched, hence indicating that it could be picked up. This was done so that consistency would be maintained with the change of colour of pieces in the main experimental task. The subjects were told that objects in Virtual Environments are not solid and are hence penetrable by the virtual hand. They were also made aware of the fact that the virtual environment does not simulate gravity. Finally, the experimenter guided each one of the subjects through this familiarisation process by talking to them and by making himself clearly present in the real world.

This was a practice that was avoided during the main experimental task, since talking to subjects during their VR experience is likely to adversely affect the sense of presence. During the main experiment, assistance and guidance was given only in cases of emergency (e.g. when a subject was in danger of colliding with an object in the real world), or in cases where the subject had specifically requested some help.

As a result of this training session, the subjects were expected to be familiar enough with the VR controls so as to be able to operate efficiently in the main experimental Virtual Environment. Finally, it should be noted that subjects who were in the non-immersed subject group for the main experiment, were also non-immersed during this VR experience.

(f) The Virtual Tri-Dimensional Chess Task and the Reproduction Session

In these sessions the subjects had to carry out the tasks as described in section 3.2(c). All relevant instructions were given in written form to the subjects and were also verbally repeated by the administrator. For example,

"Your task is to remember the new positions of the pieces on the board. You may take as long as you want to look at the board, until you feel confident that you remember the new positions. If you feel unsure, you can repeat the process, by pressing the red button, as many times as you wish."

It was made clear to them that they would have as much time as possible at their disposal, and that it was important that they felt confident that they would be able to reproduce the moves at the end of the VR session. They were also reminded of the fact that their administrator would not interact with them unless specifically asked to do so.

Finally, during the task reproduction session the administrator took clear and precise notes of the moves performed by each subject. These notes were the main source of experimental data. Also, the times between button presses were internally recorded in relevant files.

(g) The Post-Questionnaire

The post-questionnaire was given to subjects at the end of all the sessions. This included questions on their confidence about their performance, nausea caused by the VR, and three questions relating to presence. These were the same three questions that we have used in a majority of our previous studies, and recorded on a 1 to 7 scale.

- the sense of "being there" in the environment depicted by the display;
- the extent to which there were times that the virtual world seemed more the presenting reality than the real world, and

- the sense of having visited a place rather than seeing images.

3.5 Variables Measured

(a) Response (Dependent) Variables

The major response variable was the number of correct moves (out of 7 or 9) that the subject made using the real TDC. We denote this variable by C.

Presence was a dependent variable in one analysis, and an explanatory variable in another. We used the same measures of reported presence as in our previous studies. The presence variable (p) was taken as the number of 6 or 7 answers to the three questions as stated in 3.4(g), and hence was a count out of 3. As before (Slater, Usoh and Steed, 1995) an alternative score was constructed by combining the three presence question scores into a single scale using principal components analysis (Kendall, 1975). The first principal component is the linear combination of the original variables that maximises the total variance. The second is orthogonal to the first and maximises the total residual variance. The first two principal components accounted for 83% of the total variation in the original three variables. The single presence score was taken as the norm of the vector given by the first two principal components.

(b) Independent Variables

These were given by the experimental design as Immersion: Egocentric or Exocentric, Environment: Plain or Garden, and Number of Moves: 7 or 9, as discussed in Section 3.2.

(c) Explanatory Variables

These were recorded from the questionnaires and during the experiment. The major ones are recorded in Table 2. The most important ones are given earlier in the table. The "practice" variable was recorded in order to allow for different natural learning times amongst the subjects. "Remember" and "capable" although not used directly in the analysis were useful as cross checks (correlating with "practice") to check whether subjects saw sufficient practice sessions according to the needs of their level of confidence and memory.

Table 2

The Main Explanatory Variables

Variable	Brief Description
Practise (P)	The number of practice sequences initiated by the subjects.
Gender	Male (1), Female (2).
Spatial (SAT)	Results of spatial awareness test. Higher score means "better" spatial ability.
Chess	Whether the subject knows how to play chess: Yes (1), No(2).
Computer literacy	Whether or not the subject is a regular computer user: Yes(1), No(2).
Time	The overall viewing time of the moves in the VE.
Sick	Level of sickness as a result of the VE experience (1-7 scale).
Remember	The confidence with which the moves were remembered (1-7 scale).
Capable	The confidence that they had correctly reproduced the moves (1-7 scale).
Age	Age of subject in years

4. Results

4.1 Statistical Analysis for Immersion and Presence

Here the issue is the relationship between presence and the two main independent variables, immersion and environment. The dependent variable (p) was taken as the number of 6 or 7 answers to the three questions as stated above. This situation may be treated by logistic regression (Cox, 1970), where the dependent variable is binomially distributed ("number of successes out of 3 trials"), with expected value related by the logistic function to a linear predictor.

Immersion was significant at the 5% level as an independent variable for p, which was significantly higher for the egocentric compared to exocentric case. (The change in deviance was 5.623, which should be compared with a $\chi 2$ deviate on 1 d.f. = 3.841). However, the environment variable was not significant. When the analysis was repeated using the principal components score for presence, the inclusion of immersion was similarly significant at the 5% level provided that the spatial ability test score was included in the model. Interestingly, there was a small but statistically significant negative association between presence and the SAT score for those in the egocentric group.

4.2 Statistical Analysis for Performance

Table 3

Mean and Standard Deviations of Proportions of Correct Moves

Immersion:	Exocentric	Egocentric
Environment:		
Plain	0.50 ± 0.26	0.80 ± 0.22
Garden	0.61 ± 0.39	0.93 ± 0.12

Performance was measured by the variable "Correct (C)", the number of correct moves made by subjects out of 7 or 9. The mean proportion of correct moves was 0.70±0.30. There were no significant differences in number of correct moves regarding those subjects who were given 7 or 9 moves to remember. Table 3 shows the means and standard deviations for the proportions of correct moves. This suggests that the task performance improves with egocentric compared with exocentric, and with the richer environment. However, this does not take into account the influence of other possible factors, so we use logistic regression for more thorough analysis.

Table 4

Logistic Regression of Number of Correct Moves

Overall fit: Deviance = 17.626 on 16 d.f

Variable	Parameter Estimate	Standard Error	Change in Deviance
Overall mean	0.9535	2.022	
Egocentric Immersion	1.127	0.4538	6.61
Garden Environment	1.820	0.5420	12.41
No Previous Chess	-2.477	0.8459	9.36
Practise	0.5608	0.2292	9.16
Female	-13.47	4.476	11.84
SAT (female)	0.1909	0.06631	11.15

Now treating C ("correct") as a response variable, we may consider it as a binomially distributed variable being the number of correct moves out of n (= 7 or 9) possibilities, and use the logistic regression model outlined in Appendix A. The null hypothesis is equivalent to the subjects simply guessing moves at random, rather than based on their gained understanding of the spatial layout and the moves themselves. The independent variables (immersion, environment) and each of the explanatory variables of Table 2 were considered in the analysis.

The results are shown in Table 4, and the null hypothesis is rejected. For a good fit of the data to the logistic regression model, the overall deviance should be small, so that a value of less than the tabulated value is significant. Indeed the overall deviance is approximately equal to the degrees of freedom, which is what is expected for a good fit. No term can be deleted from the model without significantly increasing the deviance. This is shown in the last column of the table. The values are the increases in deviance that would occur were the corresponding term to be deleted from the model. These should be compared with the tabulated $\chi 2$ deviate on 1 d.f. = 3.841. Further analysis is presented in Appendix B.

4.3 Interpretation of Results

The results suggest the following - that other things being equal:

• Performance as measured in this particular experiment is positively associated with egocentric immersion in comparison with the exocentric screen based viewpoint;

• Performance is positively associated with a more realistic (garden) environment compared with an empty environment;

• Females do not do as well as males in this particular experiment, but

• In the case of females a higher SAT score is associated with increased performance, whereas for males the SAT score is not associated with performance.

These results also take into account that (other things being equal):

• Previous knowledge of chess is associated with better performance;

• Better performance is achieved the greater the number of practice sequences.

5. Conclusions

In this paper we have distinguished between immersion and presence, and considered the relationship of these to performance. We have argued that immersion is a description of a technology, whereas presence is concerned with the concomitant behavioural and psychological responses of people. In discussions of "performance" in relation to VEs it should always be clear as to whether the performance relates to efficiency regarding the performance of some task within the VE or in the real world subsequent to a VE experience. We argue that although increased immersion may well improve performance in certain tasks due to the higher quality and quantity of information available, there is no particular reason to expect presence to improve performance. Presence is concerned with how well a person's behaviour in the VE matches their behaviour in similar circumstances in real life, rather than with how well they perform as such.

We have carried out a case-control experiment to study the relation between immersion and performance for a task involving comprehension and memory of a complex 3D object, events in relation to that object, and the subsequent reproduction of those events in the real world. The results suggest that increased immersion (egocentric rather than exocentric viewpoint, and greater vividness in terms of richness of the portrayed environment) do indeed improve task performance. The results take into account relevant background knowledge (chess experiences) and possible gender and spatial ability differences. They also take into account possible differences in learning speed (practice). It is sometimes suggested that females are less good at spatial reasoning than males. This study suggests that the better females are at spatial reasoning, the better their performance in this experiment, whereas spatial ability as measured by the SAT was not correlated with improved performance for males.

The study also found that reported presence was higher for egocentric compared to exocentric immersion, but that presence itself was not associated with task performance.

As always such experiments raise more questions than they answer. We noted that for both our independent variables there was no clear cut distinction in levels of immersion. Although the egocentric viewpoint was "surrounding" in terms of our definition of immersion, it also had a lower resolution (and thus was less "vivid"). Although the garden environment was more vivid, it was also associated with a lower frame rate. The degree of immersion with respect to the "surrounding" component is clearly not binary - we have considered two extremes. What would happen with larger and larger screens as a means of increasing immersion? How would our experiment fare in the CAVE? What would have happened had both environments been the same but one texture mapped and the other just Gouraud shaded? What would have happened had their been a longer delay before asking subjects to reproduce the moves? It is extremely challenging to set up an experiment (with limited resources) that can be clear cut. But this is probably evidence that research into virtual environments remains in its infancy.

Appendix

Analysis of Correct Moves

From the information in Table 4 we can construct the linear predictor for number of correct moves for any particular combination of the factors and variables shown. For example, suppose we require the estimated linear predictor (η) for the case of exocentric immersion, plain environment, no previous chess experience, and male. Then

$\eta = 0.9535 - 2.477 + 0.5608 \times \text{practice}$

$\quad = -1.5235 + 0.5608 \times \text{practice}.$

Note that where an item is not shown in Table 4 (for example Male) then its parameter estimate is zero. For the items that are variables (practice, SAT) the parameter estimates are coefficients, whereas for the factors (immersion, environment, gender, chess) the coefficients are either present or zero.

Now consider the case of egocentric immersion, garden environment, previous chess experience, and female. Then

$\eta = 0.9535 + 1.127 + 1.820 + 0.5608 \times \text{practice}$

$\quad\quad -13.47 + 0.1909 \times \text{SAT}$

$= -9.570 + 0.5608 \times \text{practice} + 0.1909 \times \text{SAT}$

Body Centred Interaction

"Well then, what about the actual getting of wisdom? Is the body in the way or not...? I mean, for example, is there any truth for men in their sight and hearing? Or as poets are forever dinning into our ears, do we hear nothing and see nothing exactly?" (Socrates, Phaedo, 65A).[1]

1. Introduction

The technology to immerse people in computer generated worlds was proposed by Sutherland in 1965, and realised in 1968 with a head-mounted display that could present a user with a stereoscopic 3-dimensional view slaved to a sensing device tracking the user's head movements (Sutherland 1965; 1968). The views presented at that time were simple wire frame models. The advance of computer graphics knowledge and technology, itself tied to the enormous increase in processing power and decrease in cost, together with the development of relatively efficient and unobtrusive sensing devices, has led to the emergence of participatory immersive virtual environments, commonly referred to as "virtual reality" (VR) (Fisher 1982; Fisher et. al. 1986; Teitel 1990; see also SIGGRAPH Panel Proceedings 1989,1990).

Ellis defines virtualisation as "the process by which a human viewer interprets a patterned sensory impression to be an extended object in an environment other than that in which it physically exists" (Ellis, 1991). In this defini-

---

1.Socrates: Great Dialogues of Plato, translated by W.H.D Rouse, A Mentor Classic, 1956,

tion the idea is taken from geometric optics, where the concept of a "virtual image" is precisely defined, and is well understood. In the context of virtual reality the "patterned sensory impressions" are generated to the human senses through visual, auditory, tactile and kinesthetic displays, though systems that effectively present information in all such sensory modalities do not exist at present. Ellis further distinguishes between a virtual space, image and environment. An example of the first is a flat surface on which an image is rendered. Perspective depth cues, texture gradients, occlusion, and other similar aspects of the image lead to an observer perceiving three dimensional objects. The second, a virtual image, is the perception of an object in depth, leading to accommodation, convergence, and possibly stereopsis - for example, as might be generated by a pair of binocularly separated pictures fused to provide a stereoscopic image. The third, a virtual environment, incorporates the observer as part of the environment, so that head motions result in motion parallax from the observer's viewpoint, and a number of physiological and vestibular responses associated with focusing and object tracking are stimulated.

The human participant is "immersed" in the virtual environment (VE) in two ways. First, through the VE system displaying the sensory data depicting his or her surroundings. Part of the immediate surroundings consist of a representation of the participant's body and the environment is displayed from the unique position and orientation defined by the place of the participant's viewpoint within the environment. (We mean "display" and "viewpoint" with respect to all sensory modalities). Body tracking devices, such as electromagnetic sensors enable movements of the person's whole body and limbs to become part of the dynamic changes to objects in the VE under his or her immediate control (see Kalawsky, 1993). This is the second aspect of immersion: that proprioceptive signals about the disposition and dynamic behaviour of the human body and its parts become overlaid with consistent sensory data about the representation of the human body, the "Virtual Body" (VB). Putting this another way: proprioception results in the formation of an unconscious mental model of the person's body and its dynamics. This mental model must match the displayed sensory information concerning the VB. The VB is then under immediate control of the person's motor actions, and since the VB is itself part of the displayed VE, the person is immersed in the VE. We call such environments "Immersive Virtual Environments" (IVEs).

The term "immersion" is a description of a technology, which can be achieved to varying degrees. A necessary condition is Ellis' notion of a VE, maintained in at least one sensory modality (typically the visual). For example, a head-mounted display with wide field of view, and at least head tracking would be essential. The degree of immersion is increased by adding additional, and consistent modalities, greater degree of body tracking, richer body representations, decreased lag between body movements and resulting changes in sensory data, and so on.

Immersion may lead to a sense of presence. This is a psychological emergent property of an immersive system, and refers to the participant's sense of "being there" in the world created by the VE system. Note that immersion is a necessary rather than a sufficient condition for presence - immersion describes a kind of technology, and presence describes an associated state of consciousness.

In addition to the necessity of an immersive technology, the interaction techniques in a virtual reality may also play a crucial role in the determination of presence. For example, if through the limitations of body tracking, people must carry out everyday activities in an unnatural or artificial way, for example, moving through the world by

pointing, this may lever them out of the illusion provided by the VE, thus reducing the sense of presence. In this Chapter we introduce a paradigm for interaction in IVEs called "Body Centred Interaction" (BCI). The fundamental idea is that interaction techniques that maximise the match between proprioceptive and sensory data will maximise presence, within the constraints imposed by the display and tracking systems.

In the next section we examine the role of the body in everyday reality, and the VB in virtual reality. We consider presence more closely in Section 3. The BCI paradigm is examined in detail in Section 4, together with a number of examples, including walking, scaling and communication. In Section 5 we discuss the use of the VB in communication between human participants. Conclusions are presented in Section 6.

2. The Body

2.1 The Physical Body in Everyday Reality

Possession of a body is so obvious that its major functions can be overlooked (Synnott, 1993). It fulfils several crucial functions. It is:

- • The physical embodiment of self;
- • The medium of interaction, through the use of our bodies we interact with and are able to change the world;
- • The anchor of the self in the sensory world: our sensory organs receive data about external reality which our mind/brain system interprets as perceptions of the world;
- • A medium of communication: it allows us to communicate with other humans through the use of sound and gestures. By changing the world we construct powerful media of communications.
- • It is the social representation of self in several respects: we recognise the existence of others through their bodies, we decorate our bodies in various ways to indicate aspects of our social status, and so on.

The body is our connection with reality, it is the means through which we participate in everyday reality. Our sensory organs take in data about external reality which leads to perception, cognition and eventually to behaviour which converts this information into meaningful action through which we change external reality.

It is a relatively recent view that it is through the body and sensory perception that we come to understand reality. For example the ancients held the belief that the body is what prevents us from knowing reality:

Socrates:

"And I suppose it [the soul] reasons best when none of these senses disturbs it, hearing or sight, or pain, or pleasure indeed, but when it is completely by itself and says good-bye to the body, and so far as possible has no dealings with it, when it reaches out and grasps that which really is." [1]

It is a fundamental part of modern scientific, and perhaps common sense thought, that sense perceptions are the ultimate foundation of our knowledge about ourselves and the world.

2.2 Proprioception

Proprioception is defined by Oliver Sacks as "... that continuous but unconscious sensory flow from the movable parts of our body (muscles, tendons, joints), by which their position and tone and motion is continually monitored and adjusted, but in a way which is hidden from us because it is automatic and unconscious" (Sacks, 1985). Proprioception allows us to form a mental model that describes the dynamic spatial and relational disposition of our body and its parts. We know where our big left toe is, without looking, by relying on this body model. We can touch our nose with our right forefinger, with closed eyes, similarly by relying on this unconscious mental model formed from the proprioceptive data flow.

Sacks quoted the philosopher Wittgenstein in pointing out the fundamental nature of the proprioceptive sense, considered by many as a kind of hidden "sixth sense":

Wittgenstein:

"The aspect of things that are most important for us are hidden because of their simplicity and familiarity. (One is unable to notice something because it is always before one's eyes). The real foundations of his enquiry do not strike a man at all".

Proprioception is best appreciated when lost: Sacks describes the case of a woman who lost this sense, and was unable to move her body under conscious control. It was only through visual feedback, by looking in a mirror, that she was eventually able to move with conscious volition.

2.3 Virtual Bodies

Virtual reality offers a challenge to the everyday relationship between mind and body. This relationship is so fundamental that we normally do not think about it. Only in times of injury and crisis does the relationship come to the fore. However, entering into a virtual reality can be a shock: based on sensory data the mind may be fooled into the illusion of being in an alternative world - the results of head tracking strongly confirm this, since a turn of the head to the right swings the world to the left as in everyday reality. Motion parallax and stereopsis provide

---

1.Plato,The Phaedo, op. cit.

further evidence. And yet --- look for what you would expect to see - your own body, and it may be missing, perhaps replaced by a disembodied polygonized "hand".

The proprioceptive stream is informing us, as always during the conscious state, that the body is still there as usual. The sensory data contradicts this, there is no body. The virtual body concept is an attempt to reduce the contradiction between sensory data and proprioception by constructing a body representation slaved to the available tracking devices.

Our programs and experiments outlined in this Chapter were implemented on a DIVISION ProVision200 system. The ProVision system includes a DIVISION 3D mouse, and a Virtual Research Flight Helmet as the head mounted display. Polhemus sensors are used for position tracking of the head and the mouse. Scene rendering is performed using an Intel i860 microprocessor (one per eye) to create an RGB RS-170 video signal which is fed to an internal NTSC video encoder and then to the displays of the Flight Helmet. These displays (for the left and right eye) are colour LCDs with a 360×240 resolution and the HMD provides a horizontal field of view of about 75 degrees. The frame update rate achieved during the experiments was about 10-15 frames per second.

With the VB we have used throughout participants see a representation of their right hand, and their thumb and first finger activation of the 3D buttons on the DIVISION 3D mouse, are reflected in movements of their corresponding virtual finger and thumb. The hand is attached to an arm, that can be bent and twisted in response to similar movements of the real arm and wrist. The arm is connected to an entire but simple body representation, complete with legs and left arm. Forward movement is accompanied by walking motions of the virtual legs. When participants turn their real head around by more than 60 degrees, then the virtual body is reoriented accordingly. So for example, if they turn their real body around and then looked down at their virtual feet, their orientation lines up with their real body. However, turning only the head around by more than 60 degrees and looking down (an infrequent occurrence), results in the real body being out of alignment with the virtual body.

3. Presence

3.1 The Absence of Presence

An IVE may lead to a sense of presence for a participant taking part in such an experience. Presence is the psychological sense of "being there" in the environment based on the technologically founded immersive base. However, any given immersive system does not necessarily always lead presence for all people: the factors that determine presence, given immersion, is an important area of study (Barfield, 1993; Held and Durlach, 1992; Heeter, 1992; Loomis 1992a; Sheridan, 1992; Slater and Usoh, 1994a,c; Zeltzer, 1992).

Like proprioception, presence is so fundamental to our everyday existence that it is difficult to define. Imagining the loss of presence is more difficult than imagining the loss of proprioception. The concept of presence "no where" is logically unsound, since presence implies a "somewhere". Equating loss of presence with loss of con-

sciousness does not lead to any further understanding. However, it does make sense to consider the negation of a sense of presence as the loss of locality, such that "no presence" is equated with no locality, the sense of where self is as being always in flux. Interestingly, Sacks describes the case of a man without the capability for present day memory. It was essentially impossible to have a conversation with him, since the context would be lost after a few moments, when he forgot who he was talking to, and what the conversation was about. This is a kind of neurological loss of presence. Imagine a VR system that continuously and randomly changed the environment, so that the human participant could form no stable sense of locality, and no relationship with any object: everything being continually in flux. Such an environment would not be presence inducing.

3.2 Presence and the Body

It can be argued that there is an inherent logical connection between the degree of presence and the VB. If the match between proprioception and sensory data about the corresponding dynamics of the body is high, then the person immersed in the VE is likely to identify with their VB. If sensory data confirms that this VB functions effectively within the larger (computer generated) environment, then there must be presence within that environment. The VB has become identified with "self", the VB is immersed within a particular environment, therefore self must be in that environment.

There is empirical evidence from a number of case-control studies providing evidence for this idea. The first pilot study divided 17 subjects into two groups, experimental and control. The experimental group had a VB as described in Section 2, and the control group had a very impoverished VB consisting only of a 3D arrow pointer that responded correctly to (right) had movements and orientations. All subjects carried out the same tasks, which involved moving from a corridor into a number of rooms, and each room exercised a different aspect of the experiment. For example, in one room objects spontaneously flew towards the face of the subjects, and in another, they were perched on a plank over the edge of a precipice.

In this experiment presence was measured in two ways. The first was by a particular question in a questionnaire administered after the experience (To what extent did you experience a sense of being "really there" inside the virtual environment?). This was measured on a 6 point scale, from 1 = "Not at all really there" to 6 = "totally there".

The second method was to observe the reactions of the subjects to "danger" - in particular did they exhibit the looming effect when objects flew towards their faces (ie, did they "duck"), and second, did they react in an observable manner, including verbal exclamations, when over the virtual precipice. The results suggested a positive association between the VB and the observed reaction to "danger". If a reaction to danger indicates presence, then possession of a VB did positively influence presence. These results are extensively reported in (Slater and Usoh, 1992; 1993a). A first analysis did not find a positive relationship between VB and reported sense of presence as indicated by the responses to the questionnaire.

The situation was more complex than this, however. We were puzzled by the fact that these 17 people had all had very similar experiences, and yet their reactions were so different to one another, including their responses to the presence question. The human participant in a VE does not simply absorb the VE generated sensory data, but processes this through the mental models and representation systems typically employed by the person in everyday reality. Since people have different models of the world and corresponding preferences in (unconsciously) processing sensory data, and since the VE typically offers very biased sensory data (ie, very much biased towards the visual), this might explain the variation in people's responses.

We carried out a post-hoc analysis of the questionnaire data, including an analysis of essays written by the subjects twenty four hours after the end of the experiment. This was based on a neuro-linguistic programming (NLP) model of subjective experience, which states that all such experience is encoded in terms of three main representation systems, Visual, Auditory and Kinesthetic (VAK) (Dilts et. al., 1979). The Visual system includes external images and remembered and constructed internal images. The Auditory system includes external sounds, and internal remembered and constructed sounds. It also includes internal dialogue, that is the person talking to him- or her-self on the inside. The Kinesthetic system includes kinesthetic and tactile sensations and also emotional responses (which are decomposed into specific patterns of internal tactile and kinesthetic sensations). The model claims that people have a tendency to be dominant in one or other of these systems, and that such dominance may be reflected in language patterns: specifically, in the (visual, auditory, kinesthetic) predicates and references they tend to use. For example, when a person says "I see what you mean", this is taken not just as an arbitrary and accidental choice of expression, but as an indication of their internal processing - they may be literally making an internal picture of the situation under discussion. They could equally well have said "I hear what you're saying" or "I have a feeling for what you say", but instead chose the visual predicate.

NLP also distinguishes between egocentric and exocentric perceptual positions. The perceptual position is the standpoint from which the person experiences and remembers events. A person might remember an event from an associated (egocentric) standpoint, and see the event unfolding in his mind's eye from the viewpoint in which it was originally experienced. This is called the first perceptual position. Alternatively a person might remember the event from a dissociated (exocentric) perspective - either from the point of view of another actor in the scene (second position), or from an abstract, disembodied point of view (third position). For example, a person trying to convince someone in an argument might say: "I can feel that it is right" (first position, K) or "You can tell that it is right" (second position, A) or "It can be seen that it is right" (third position, V). The representation systems and perceptual position are logically orthogonal - there being nine possible combinations in this example.

Using the essays written by the subjects as part of the post-experiment information that we collected, we counted the number of V, A, K predicates and references used as a proportion of the total number of sentences written by each subject. Similarly, we classified each sentence as belonging to either the first, second or third perceptual position. Hence variables were constructed that attempted to measure the extent of dominance with respect to representation system and perceptual position for each subject in the experiment, and these were included as explanatory variables in a statistical (regression) analysis of the data with the reported degree of presence taken as the dependent variable.

Since the VR system we were using presented the participant mainly with visual information, we expected - if the NLP hypothesis were useful - that visual dominance would be positively correlated with reported presence, and auditory dominance negatively correlated. The results were rather startling - even though the regression analysis was not statistically secure (the dependent variable being a measurement on an ordinal scale) the explanatory power of the model was very high indeed, with a multiple squared correlation coefficient of 0.99, and with a very high level of fit (better than 1% significance). The regression model resulted in the following conclusions:

(a) That independently of whether or not the subject has a virtual body, the higher the proportion of visual predicates and references used, the greater the sense of presence, and the higher the proportion of auditory predicates and references the lower the sense of presence.

(b) For those with a virtual body, the higher the proportion of kinesthetic references and predicates the higher the sense of presence. For those without a virtual body, the higher the sense of kinesthetic terms the lower the sense of presence.

(c) The level of presence increases with first perceptual position (P1) up to the mean level of P1, and then decreases. (The model was quadratic in P1). This is the same for each group, except that the rate of change is steeper for those in the control group.

The analysis and results are reported in (Slater and Usoh, 1993a; 1994a,c).

It is result (b) that is most interesting in the present discussion. It indicates a relationship between kinesthetic dominance, the VB and reported degree of presence. The K system is the system of the body - it is very strongly related with proprioception as discussed in Section 1. This result gave us a clue that there is a relationship between the VB, proprioception and presence.

The experiment described here was only a pilot, and it was unsatisfactory from the point of view of direction of causality. We could not say that representation systems were a causal factor in presence, since the data used for measuring these was obtained after the VR experience. It could have been said that that experience itself was a causal factor determining the representation systems used when writing about it. Therefore, we carried out a further major study, with 24 subjects, where we used a questionnaire to assess dominant representation systems and perceptual position well before the VR experience. This study, where each participant did have a VB, resulted again in a model with very strong explanatory power for the representation systems, but no significant effect was found for perceptual position. Again, the higher the visual dominance the greater the degree of presence, the higher the auditory dominance, the lower the degree of presence, and also (this time since all had the same VB) the higher the kinesthetic dominance, the higher the degree of presence. The experiment and results are discussed fully in (Slater, Usoh and Steed, 1994c).

This experiment used a more comprehensive measurement of presence based on:

(a) The subject's sense of "being there" - a direct attempt to record the overall psychological state with respect to an environment;

(b) The extent to which, while immersed in the VE, it becomes more "real or present" than everyday reality;

(c) The "locality", that is the extent to which the VE is thought of as a "place" that was visited rather than just as a set of images.

This last is similar to the idea of Barfield and Weghorst who write that "... presence in a virtual environment necessitates a belief that the participant no longer inhabits the physical space but now occupies the computer generated virtual environment as a 'place'" (op. cit., p702). Each of these was measured on a 7 point scale, and the overall score for an individual was the number of highest scores (6 or 7) out of three.

Especially interesting in this experiment is that we programmed the virtual left arm and hand to mirror the movements of the corresponding right hand limbs. The idea was to see the extent to which subjects would match their real left hand with the virtual one. Four out of the 24 subjects exhibited this matching behaviour. These four subjects had a significantly higher score on the K representation system than the other subjects (in fact by more than double). We speculate that these subjects had a requirement to match the proprioceptive with the sensory data. They saw their virtual left hand move, and the only way that the matching was possible was to move their real left hand in conjunction.

These four subjects must have had a very high degree of identification with their virtual bodies. In our first pilot experiment, where the virtual left arm was in a fixed position, some of the subjects wrote about their confusion or perhaps lack of identification with the VB. Strange effects were observed, and recorded:

• One subject on noticing the fixed virtual left arm began to move her real left arm very rapidly, in a manner indicating panic.
• Another wrote "I thought there was really something wrong with my [left] arm";
• Others talked of their virtual bodies being - "a dead weight", a useless thing", "nothing to do with me".

Such remarks were reminiscent of Sack's patients who lost the proprioceptive sense in some of their limbs. This suggests that the lack of a normal relationship between the proprioceptive system and the behaviour of the VB could be very important factor in people's acceptance of and responses to immersive virtual environments.

3.3 Presence Summary

In this Section we have examined the concept of presence in a VE, and in particular the relationship between the physical body, virtual body and presence. There are three aspects to the relationship that we have discussed so far. The first is that proprioception provides a sense of the physical body and its activities, leading to a mental body model. Presence is likely to be enhanced the more that this mental body model behaviourally matches the virtual body representation in the VE. Since the participant is only aware of this VB through the sensory (mainly visual) data supplied by the immersive system, presence requires that proprioceptive data be continually overlaid with consistent virtual sensory data. The second, is that evidence suggests that, other things being equal, a virtual body will, in any case, enhance the sense of presence. Third, the body is the repository of the sensory apparatus, which in turn leads to the fundamental representation systems based on the senses (visual, auditory and kinesthetic). The representation systems are a powerful factor in explaining people's reported sense of presence. In particular, this is true for people who are dominant on the kinesthetic representation system - that is, those for whom proprioceptive data (how they "feel") is an important explicit and verbalised component of their mental processing.

The unique feature of modern virtual reality systems is that they are general purpose presence transforming machines. Systems and applications have existed for many years that provide a high degree of presence: flight simulators are an obvious example. However, such systems always provide a very high sense of presence within a particular and fixed environment. A flight simulator can, for example, never be used to provide a sense of presence within a supermarket. An IVE system, can, however, be used to provide a sense of presence in an airplane cockpit, and also in a supermarket: it is only a question of the database and interaction model used. Obviously, since a flight simulator is specialised to airplanes it is typically much more successful than a virtual reality system for its particular application domain: but at the great cost always associated with very special purpose systems. The choice between an IVE and a traditional simulator then becomes a question of economics.

Steuer has gone as far as taking presence as the defining feature of VR: "A virtual reality is defined as a real or simulated environment in which a perceiver experiences telepresence" (Steuer, 1992). We are tempted to extend this definition to include the importance of the VB:

> A virtual reality is a real or simulated reality in which the self has a (suspension of dis-) belief that he or she is in an environment other than that which his/her real body is located. Self perceives sensory information correlated with proprioceptively valid feedback about the behaviour and state of his/her body in that environment.

We have concentrated here on presence as the central phenomenon of virtual reality, and have examined its relationship to the body and VB. In the next Section we show how we have exploited these relationships in the construction of interactive techniques.

4. Body Centred Interaction

4.1 Motivation and Concepts

In the first pilot experiment on presence discussed in Section 3, we observed that some subjects found it exceedingly difficult to move around the VE using a navigation metaphor based on hand gesture pointing. For example, the following are reports from their essays written after the experience:

> "Sometimes [I had] a desperate need to actually walk when virtually walking, there does seem to be a conflict between what the eyes see and the body feels - eg, my feet appear to be floating but I can feel my feet on the ground."

> "Trying to separate virtual and physical movement: constantly being aware - my initial response was to make the physical move then forcing myself to use the mouse instead... The amount of concentration I had to use was something I remember particularly. Moving around with the mouse, forwards and backwards - and with the helmet turning around - it was difficult to reconcile the two ways of moving."

This illustrates in the negative the central idea of the BCI paradigm: interaction techniques should be constructed so that there is a match between sensory data ("what the eyes see") and proprioceptive feedback ("what the body feels"). The typical approach is to either overload almost all forms of interaction onto a set of hand gestures or manipulations (Vaananen and Bohm, 1993; Brooks et. al., 1990) or to use inappropriate methodology taken from screen based interfaces, such as menus and icons. We are reminded of a famous passage written by Marx:

K. Marx:

> "Men make their own history, but they do not make it just as they please; they do not make it under circumstances chosen by themselves, but under circumstances directly encountered, given and transmitted from the past. The tradition of all the dead generations weighs like a nightmare on the brain of the living. And just when they seem engaged in revolutionising themselves and things, in creating something that has never yet existed ... they anxiously conjure up the spirits of the past to their service and borrow from them names, battle cries and costumes ..."[1]

Virtual reality must, on the contrary, invent its own new ways of thinking, appropriate and native to the new technology.

Body Centred Interaction involves a number of components:

(a) Inference about the state of the body from limited information

---

1.K. Marx, The Eighteenth Brumaire of Louis Napoleon, in Marx and Engels, Selected Works in One Volume, Lawrence and Wishart Ltd, 1968.

One of the concepts of the BCI approach is the construction of an abstract (device independent) control model that defines the mapping between physical tracking capabilities and the associations with virtual body dynamics. For example, consider two extremes - a full body suit that tracks the position of all the major limbs of the body, compared to a six degrees of freedom 3D mouse held in one hand. It is assumed in both cases that there is a HMD that tracks the position and orientation of the head. Now in the former case, there is a relatively straightforward mapping between the tracking information and the position and orientation of the virtual body and its limbs. In the latter case, only the head position and orientation and the position and orientation of one hand is known. Hence in this case, the position and orientation of the VB as a whole is a matter for inference. The objective is to construct a consistent inferential model for this mapping. The discussion in Section 2.3 illustrates a primitive example of this.

(b) Body centred feedback

Interaction requires feedback about the state of the VB, and its relationship to the environment. This involves the generation of real-time shadows and reflections, that include the VB (as well as shadows of objects generally). It also involves the use of a graphics viewing model that simulates and stimulates peripheral vision, in spite of the relatively small field of view actually provided by the visual display devices.

In previous work (Chrysanthou and Slater, 1992) we have constructed an algorithm for dynamic shadows in the context of polygonal scenes illuminated with local lighting. Shadows are well-known to be important in understanding spatial relationships (Puerta, 1989). The shadow of the person's own VB would be an exciting method for feedback in this context. Mirrors and reflections are an obvious extension of this work.

Today's HMDs typically provide a reduced field of view compared to the average human FOV. Hence, unlike the situation in everyday reality, the participant is typically not always aware of the state of his virtual body, or of events that would normally be signalled by peripheral vision. We have developed a graphics viewing pipeline that does simulate peripheral vision, and have shown experimentally that it is possible to stimulate the behaviour associated with peripheral vision in spite of the relatively small FOV of HMDs (Slater and Usoh, 1993b).

We are currently developing implementations of both the rapid shadow and peripheral vision models on the VR system.

(c) Magical and Mundane Interaction

Interaction is the ability of the participant to move through and change the world, that is, navigation and manipulation. This falls into two further categories, which we call *mundane* and *magical*. Mundane interaction is that which attempts to faithfully reproduce a corresponding interaction in everyday reality. For example, the process of picking up an object, or driving an automobile. Magical interaction involves actions that are not possible in everyday reality - such as a person flying by his or her own volition, walking through walls, tele-portation - that is moving instantaneously from place to place, psycho-kinesis - that is, action on an object at a distance, and other similar examples. Table 1 classifies these types of interaction.

Table 1

Magical and Mundane Interactions

Interaction	Examples	Manipulation Examples	Navigation Examples
Mundane Reproduction of interactions from the world of everyday reality.	picking something up; walking; driving an automobile.	object selection and placement; transformations, deformations.	walking; driving or flying a vehicle; space walks.
Magical Production of interactions that are only imaginable in everyday reality.	flying by own volition; tele-portation; psycho-kinesis.	scaling the environment; psycho-kinesis	flying under own volition; teleportation

To the extent that a VR system is to be used as a simulation of everyday reality, for example, for the purposes of training, it is necessary for the actions that a person makes in the VE to be intuitively associated with the corresponding actions that they would need to take in everyday reality. It is also possible for magical interaction to be accomplished in an intuitive way, involving the marshalling of mental models for activities on the part of the participant that even though achieving magical effects, can seem to be accomplished naturally. We have found that interactions based directly on the use of the person's VB seem to satisfy this criterion. The following sections consider examples from both categories: mundane - walking, climbing and descending steps and ladders; magical - scaling the environment and remote object selection.

4.2 Walking: The Virtual Treadmill

A standard solution for navigation in IVEs is to make use of the hand-held pointing device. VPL used the Data-Glove (Fisher 1986; Foley, 1987) with which a hand gesture would initiate movement, and the direction of movement would be controlled by the pointing direction. Velocity was controlled as part of the gesture: for example the smaller the angle between thumb and first finger the greater the velocity.

DIVISION's ProVision system typically employs a 3D mouse (though it supports gloves as well). Here the direction of movement is determined by gaze, and movement is caused when the user presses a button on the mouse. There are two speeds of travel controlled by a combination of button presses. Other methods of navigation are discussed in (Brooks, 1992; Fairchild, 1993; Iwata and Matsuda, 1992; Mackinlay et. al., 1990; Robinett and Holloway, 1992; Song and Norman, 1993).

In the experiments mentioned above we adjusted the ProVision's standard interface, and based direction of movement on the pointing direction of the 3D Mouse. This disassociation of gaze and direction of movement gives the participant an extra degree of freedom in exploring the VE.

We mentioned above the difficulty that some subjects have using a pointing device for navigation. In some contexts such an approach might be natural, for example in a simulated space walk - but then the normal methods of moving around, such as taking one or two small steps would need to be disabled with perhaps the participant seated in a chair. The pointing method would be the only method for movement over large or small distances, so that the conflict mentioned by the subjects could not occur.

Brooks noted that "Physical motion powerfully aids the illusion of presence, and actual walking enables one to feel kinesthetically how large spaces are..." (Brooks, 1992). As part of the Building Walkthrough project at the University of North Carolina, a steerable treadmill was constructed, that allowed users to actually experience walking through virtual buildings and building sites. The Virtual Treadmill is a similar idea, but implemented only in software, and without the restrictions necessitated by a real treadmill where the user cannot step off from it in order to really walk a few steps.

The idea of the Virtual Treadmill is straightforward - whenever participants carry out the activity of walking on the spot, that is standing in one place but with leg motions similar to walking, the system moves them forward in the virtual space, with direction of movement governed by gaze. This is achieved by passing all HMD data through a pattern recogniser filter which is able to distinguish head movements characteristic of such "walking on the spot" behaviour from any other behaviour at all. Therefore, virtual ground is covered in this technique by almost really walking, or by taking one or two actual physical steps: each case involving whole body movements similar to those of walking in everyday reality. Contrast this with the usual method used in VR, which is sometimes moving by actually walking, and other times using a pointing hand gesture. In the new method there is no use made at all of the hand-held pointing device. This can be reserved solely for other forms of interaction such as object manipulation.

Two studies with users were carried out regarding the influence of the Virtual Treadmill on navigation and presence. In each study there were 16 subjects divided into experimental and control groups - the experimental group were "walkers" - they used the Virtual Treadmill idea, and the controls were "pointers" - they used the hand gesture with the 3D mouse as usual. A full report of the first study is given in (Slater, Steed and Usoh, 1993c). We concentrate here only on the results relating to presence. The task of both groups was to navigate through a room containing many obstacles, pick up an object, take it out into a corridor, and then locate and enter another room at the far side of the corridor. The objective was to place the object on a chair in that room. This chair was reachable

only by crossing a chasm over a precipice. The control group first carried out this task as "pointers", answered a questionnaire, and then repeated the experiment as "walkers", and completed a second questionnaire. The experimental group did this in the opposite order. At the end of the first part of the experiment, each group had experienced only one type of navigation technique, only "walking" or "pointing". After the second part of the experiment, each person had experienced both types. Three control group subjects were not included in the comparative part of the study because the walking technique did not work for them at all. Overall though, the pattern recogniser correctly predicted behaviour, that is it distinguished between walking on the spot and other activities with a success rate of between 85% and 95%.

Table 2 shows the results of this experiment in regard to subjective reporting on presence. There is no difference in presence between the two groups immediately after the end of Part I of the experiment, that is after each subject had experienced one method of walking. However, in the comparison after Part II, amongst those who had a preference, the walking method led to a higher subjective sense of presence. However, comparisons such as these are suspect, since it cannot be known whether the experience of the first session influenced the results of the second session.

Table 2

Subjective Reporting on Presence

Being there			Real or present			Seeing/visiting		
Please rate your sense of being there in the computer generated world...			To what extent were there times during the experience when the computer generated world became the "reality" for you, and you almost forgot about the "real world" outside?			When you think back about your experience, do you think of the computer generated world more as something that you saw, or more as somewhere that you visited?		
In the computer generated world I had a sense of "being there"...			*There were times during the experience when the computer generated world became more real or present for me compared to the "real world"...*			*The computer generated world seems to me to be more like...*		
1. not at all			1. at no time			1. something that I saw		
...			...			...		
7. very much			7. almost all of the time			7. somewhere that I visited		
Group	Mean	Median		Mean	Median		Mean	Median
Exp.	6	6		5	5		5	5

Control	5	5		4	3		5	4
Part II comparison:			**Part II comparison:**			**Part II comparison:**		
prefer:			**prefer**:			**prefer**:		
walking: 6			walking: 7			walking: 7		
same: 5			same: 5			same: 6		
mouse: 2			mouse: 1			mouse: 0		
TOTAL:13			TOTAL:13			TOTAL:13		

In the second study the scenario was slightly different. The task was to pick up an object located in a corridor, take it into a room and place it on a particular chair. The chair was placed in such a way that the subjects had to cross a chasm over another room about 20 feet below, in order to reach it. The subjects could get to the chair either by going out of their way to walk around a wide ledge around the edges of the room, or by directly moving the shorter distance across the chasm. This was a simple virtual version of the famous visual cliff experiment by E.J. Gibson (Gibson and Walk, 1960). All subjects were watched by an observer, who in particular recorded whether or not they moved to the chair by walking around the ledge at the side of the room, or by walking directly across the precipice. In the event, only four subjects out of the sixteen (two from each group) walked across the precipice.

The main conclusion from the statistical analysis was that for the "walkers", the greater their association with the VB the higher the presence score, whereas for the "pointers" there was no correlation between VB association and the presence score. Other statistically significant factors were:

> (i) path taken to the chair: a path directly over the precipice was associated with lower presence. This is as would be expected, and is useful in corroborating the veracity of the presence score.

> (ii) degree of nausea: a higher level of reported nausea was associated with a higher degree of presence. This same result has been found in each of our studies. We speculate that the sense of motion in VR is a cause of both simulator sickness and an influence on presence (McCauley and Sharkey, 1993). Finding nausea and presence associated would therefore be expected, even though there may not be a direct causal link between them. There is the further point that presence is concerned with the effect of the environment on the individual. A person who experiences nausea as a result of the VR has certainly been influenced by it!

These results were obtained from a logistic regression analysis, that is, counting the number of 6 or 7 scores across the three presence questions and using this count out of three as the dependant variable. Here the dependent variable is binomially distributed, with expected value related by the logistic function to a linear combination of independent and explanatory variables (Cox, 1970).

An alternative analysis of the same data was carried out, where the three presence scores were combined into one overall score using a principal components analysis. A statistically significant normal regression model was

obtained, with qualitatively similar results to the first analysis. The overall regression was significant at 5% with a multiple squared correlation coefficient of 0.81. Here though, instead of path to the chair being significant, a variable representing the comparison between vertigo experienced in the virtual world with what might have been experienced in the real world in a similar situation, was significant instead. Subjects were asked to rate their reaction to the visual cliff regarding the extent to which it was the same or different to what they would have expected it to be in real life. In the analysis a higher degree of presence was associated with the comparison resulting in a "same as real life". Loomis suggests that one objective way of assessing presence is the degree to which reactions are the same in virtual as in real environments (Loomis, 1992b). Again this lends support to the measure of presence used actually bearing a strong relationship to the phenomenon of presence.

This experiment, in including the degree of subjective association with the virtual body, allowed for a more sophisticated analysis. The central thesis of the BCI paradigm, that presence is likely to be enhanced with interaction techniques that attempt to match proprioception and sensory data, especially that regarding the VB, seems to be supported - since only for the "walkers" was there a positive correlation between VB association and presence. This experiment is reported in (Slater, Steed and Usoh, 1994b).

4.3 Steps and Ladders

The Virtual Treadmill has easy adaptation to other forms of navigation beside walking at ground level. Applications such as architectural walkthrough, or training for fire fighting, require participants to walk up steps or climb and descend ladders. Again, it is certainly possible to use a hand gesture, or allow participants to fly, and in some applications this would be acceptable if a degree of realism in these activities were not required. In the fire fighting example though, trainees would typically be required to carry objects (buckets, hoses, etc) while climbing steps or ladders, so that the use of hand based gestures for navigation would not be suitable. Also, in a real fire fighting situation, the fire fighters do expend energy in moving through the scenario, and here what may be thought of as a disadvantage of the Virtual Treadmill - it certainly requires more energy to perform than pressing a button or making a hand gesture - becomes an advantage in terms of realism.

At the time of writing we have adapted the Virtual Treadmill to steps and ladders in a straightforward manner. When the process monitoring collision detection notifies the system of a collision between the VB and the bottom or top rung of a staircase or ladder, subsequent walking on the spot motions will move the participant up or down as appropriate. For steps, we do not currently support walking backwards down steps (this is never a good idea in reality). For ladders, we extend the whole body gesture so that while the hand is above the head and the person is moving on a ladder, they will climb up the ladder, and while the person's hand is below their head, they will move down the ladder.

Plates 1, 2 and 3 show exterior views of a VB as it is climbing or descending steps and ladders, in one case holding a bucket.

4.4 Scaling the Environment

Scaling the environment as a whole is useful in applications where an overview of the entire scene is required, or alternatively when details need to be enlarged. This could be accomplished by defined hand gestures, or by menus and sliders. The BCI approach, however, requires the participant to carry out a whole body gesture which is semantically appropriate for the activity. Scaling the environment up is equivalent to shrinking the participant's VB. This can be accomplished by the person pushing down on his or her head with his hand and flexing the knees to lower the head, in an attempt to become smaller. Corresponding with this activity, the VB will become smaller, and the world will appear to grow larger, while the hand remains on top of the head. Shrinking the world is equivalent to growing the body. This can be accomplished with a placement of the hand under the chin, in a gesture of pushing upwards which grows the VB, and correspondingly the world appears to shrink.

This technique also supports magical navigation. Isaac Asimov's *Fantastic Voyage* can be accomplished in VR by shrinking the body to a tiny size in relation to the environment, so that the participant can move through what would in reality be microscopic spaces. (In the famous book, a doctor entered into the blood stream of a patient). Another application, would be to grow the body to a very large size, so that one small step would take the participant across to the other side of the environment. VR allows us to become microscopic creatures, or giants. The BCI paradigm tries to accomplish these magical techniques in an intuitive manner.

4.5 Body Centred Interaction Summary

The BCI paradigm therefore attempts to match sensory and proprioceptive data. An aspect of this is that it uses whole body gestures rather than limited hand based gestures or screen based interfaces in order to accomplish interactions. The goal is always to provide a gesture which corresponds in a semantic sense to the type of interaction. Hence walking is carried out by "almost walking", shrinking the body is accomplished by pushing down on the head. Other examples are easy to construct - for example selection of a distant object might be carried out by stretching the hand as far as possible away from the body. When the VR system detects such an event, it will grow the arm in the direction of pointing. Obviously, the kind of gestures possible are limited by the body tracking data available: the more of the body that is tracked, the more sophisticated can the gestures be. However, even with just the HMD tracker and glove or hand-held 3D mouse, quite a large number of different, intuitively appealing whole body gestures can be defined.

5. Communications

So far we have concentrated on a single isolated self and body within the VE. In this Section we briefly consider the implications of the BCI paradigm for people communicating in a shared VE. In this context the body becomes a social as well as a personal object. The body is not only a private representation of self, and a means

for interaction, but also a medium of communication with others. Others are represented to self through their bodies and the relationship of the body of others to that of self is extremely significant personally, socially, and culturally. In a recent book on the sociology of the body, Anthony Synnot discusses this aspect of the physical body:

Anthony Synnot (Synnot, 1993):

> "The body social is many things: the prime symbol of self, but also of the society; it is something we have, yet also what we are; it is both subject and object at the same time; it is individual and personal, as unique as a fingerprint or odour-plume, yet it is also common to all humanity ... The body is both an individual creation, physically and phenomenologically, and a cultural product; it is personal, and also state property."

Virtual bodies play a vital role in shared environments. The MultiG project at the Swedish Institute of Computer Science (Fahlen, 1992; 1993) has constructed a distributed VE where participants at physically different locations take part in, for example, joint virtual meetings. People become aware of each other in the VE through a complex function of their aura ("a space that can be seen as the enabler of interactions with other objects"), focus (a "space within which the object directs its attention") and nimbus (a space "where the object projects some aspect of its presence to be perceived by other objects"). Participants are represented by a simple VB model (a block with eyes) which is nevertheless quite powerful in representing the presence of another being.

The body in MultiG is a static entity, with no limbs. However, in meetings body posture by itself can indicate the real events which are taking place, as opposed to the superficial events at the level of verbal discussion. Body posture can be conveyed with very little information - for example, in Figure 1(a), the person depicted does not have to say anything for the observer to know what is being expressed.

Synnot shows that the face is the most powerful social symbol of self. Again, in meetings, where facial expression contradicts verbal agreement - which is likely to be more important?

Support for this kind of "body centred interaction" requires a different form of tracking technology. Rather than monitoring the body from the outside, using electromagnetic sensors such as the Polhemus, the body can be monitored from the "inside", using electrical recordings of the activity of the individual muscles or nerves, and electroencephalographic (EEG) recordings of potentials from the surface of the skull overlaying the motor cortex. There have been some applications of such biofeedback technology in VR (Lusted, Knapp and Lloyd, 1992; 1993). Such work offers great promise for a different kind of sensor and tracking technology, more in tune with the requirements of BCI.

Support for this kind of "body centred interaction" requires a different form of tracking technology. Rather than monitoring the body from the outside, using electromagnetic sensors such as the Polhemus, the body can be

monitored from the "inside", using electrical recordings of the activity of the individual muscles or nerves, and electroencephalographic (EEG) recordings of potentials from the surface of the skull overlaying the motor cortex. There have been some applications of such biofeedback technology in VR (Lusted, Knapp and Lloyd, 1992; 1993). Such work offers great promise for a different kind of sensor and tracking technology, more in tune with the requirements of BCI.

6. Conclusions

In this Chapter we have concentrated on the role of the physical and virtual body in VR. The virtual body plays a primary role in immersive virtual environments:-

- it is the representation of self;

- it is likely to be a factor in increasing presence;

- it is the foundation of a model for interaction, body centred interaction;

- it is a medium of communication with others in shared environments;

- it may lead to a theory of virtual reality, through understanding of the relationship between the physical body, the virtual body, proprioception and presence.

The essence of Virtual Reality is that we (individual, group, simultaneously, asychronously) are transported bodily to a computer generated environment. We recognise our own habitation there, through our body becoming an object in that environment. We recognise the habitation of others through the representation of their own bodies. This way of thinking can result in quite revolutionary forms of virtual communication. For example in asynchronous communication, suppose a person (X) wishes to leave a message for someone else (Y) who will enter the environment at some time after X has left. A traditional way of thinking would be to leave a written or perhaps auditory message. The VB, however, allows X to leave a copy of his or her VB there in the environment to interact with Y, to perhaps act out a scenario depicting the required information (for example, in a training application). It is these new ways of thinking that must be adopted if VR is to fulfil its potential.

Walking to the Precipice

Summary

This chapter presents an interactive technique for moving through an immersive virtual environment (or "virtual reality"). The technique is suitable for applications where locomotion is restricted to ground level. The technique is derived from the idea that presence in virtual environments may be enhanced the stronger the match between proprioceptive information from human body movements, and sensory feedback from the computer generated displays. The technique is an attempt to simulate body movements associated with walking. The participant "walks in place" to move through the virtual environment across distances greater than the physical limitations imposed by the electro-magnetic tracking devices. A neural network is used to analyse the stream of coordinates from the head-mounted display, to determine whether or not the participant is walking on the spot. Whenever it determines the walking behaviour, the participant is moved through virtual space in the direction of gaze. We discuss two experimental studies to assess the impact on presence of this method in comparison to the usual hand pointing method of navigation in virtual reality. The studies suggest that subjective rating of presence is enhanced by the walking method provided that participants subjectively associate with the virtual body provided in the environment. An application of the technique to climbing steps and ladders is also presented.

1. Introduction

The ability to get from place to place is a fundamental requirement for action in both real and virtual environments. This requirement epitomises what is very powerful, yet what also may be flawed in virtual reality (VR) systems. These offer the possibility of perceptually immersing individuals into computer generated environments, and yet the typical means for the most basic form of interaction - locomotion - do not at all match the

physical actions of walking in reality. Generally, the powerful illusion of immersion may be lost through naive interaction metaphors borrowed from non-immersive forms of human-computer interaction.

This paper describes an interactive technique for locomotion in an immersive virtual environment (or "virtual reality"). The technique is suitable in applications where the participants are constrained to ground level, for example, while exploring a virtual building, as in architectural walkthrough. The novelty of the technique is that participants carry out whole body movements in a simulation of walking, without the necessity of hardware additional to the electro-magnetic tracking devices on the head-mounted display (HMD) and glove (or 3D mouse). In brief, participants "walk in place" to move across virtual distances that are greater than the physical space determined by the range of the electro-magnetic trackers. Pattern analysis of head movements as generated by the HMD predicts whether participants are walking in place, or doing anything else at all. Whenever it is determined that they are walking in place, they are moved forward in the direction of gaze, so that the corresponding flow in the optical array gives the illusion of motion. Such illusory self-motion is usually called vection. Since the pattern analyser (ideally) only detects head movements characteristic of walking in place, participants are able to take real physical steps, while remaining within effective tracker range, without causing vection surplus to their actual movements.

In an earlier report [Slater, Steed and Usoh, 1993] we presented the technique, called the Virtual Treadmill[1], in the context of (at that time) a partially complete human-factors evaluation. In this paper we discuss the technique in the context of a model of presence in immersive virtual environments. We also present the implementation details, and results of two empirical studies with users. The utility of this idea for climbing or descending steps and ladders is also discussed.

2. Virtual Environments

2.1 The proprioceptive - sensory data loop

A virtual reality system requires that the normal proprioceptive information we unconsciously use to form a mental model of the body, be overlaid with sensory data that is supplied by computer generated displays. Proprioception was defined by Oliver Sacks as "... that continuous but unconscious sensory flow from the movable parts of our body (muscles, tendons, joints), by which their position and tone and motion is continually monitored and adjusted, but in a way which is hidden from us because it is automatic and unconscious" [Sacks, 1985]. Proprioception allows us to form a mental model that describes the dynamic spatial and relational disposition of our body and its parts. We know where our left foot is (without having to look) by tapping into this body model. We can clap our two hands together (with closed eyes) similarly by relying on this unconscious mental model formed from the proprioceptive data flow.

---

1.   The London Parallel Applications Centre has a holding patent covering the UK and other countries to protect aspects of this technology.

Tracking devices placed on the physical human body are required in order to map real body movements onto corresponding movements of the participant's self-representation in the virtual world. We call this self-representation a virtual body (VB). A fundamental requirement for an effective virtual reality is therefore that there is a consistency between proprioceptive information and sensory feedback, and in particular between the mental body model and the VB.

J.J. Gibson's notion of the ambient optical array may be employed to elaborate these ideas [Gibson, 1986]. This is conceived as an arrangement consisting of a nested hierarchy of visual solid angles all with the same apex, and completely surrounding the apex. The apex corresponds to a position in the environment, which may be occupied by an individual. Such an individual is not considered as a disembodied observer, taking up an abstract point in space, but as a live animal that continually moves through an all-surrounding environment, standing and moving on feet and with a head, eyes, ears, nose, mouth. This is not the abstract space of the mathematician.

Gibson argued that when an individual is immersed in an environment, perception of the self is inseparable from perception of the environment. When describing the occupation of a position in the ambient optical array by an individual he said that, "When the position becomes occupied, something very interesting happens to the ambient array: it contains information about the body of the observer." [Gibson, 1986, p66]. Regarding the relationship between sensory information and self perception he wrote that: "The optical information to specify the self, including the head, body, arms and hands *accompanies* the optical information to specify the environment. The two sources of information coexist." [Gibson, op. cit., p116].

This relationship between proprioceptive information and sensory data requires consistency, predictability and completeness in order to function properly. For example, when proprioceptive information arises because we have moved a leg in such a way that it comes into contact with another object - the sensory data must correctly inform us, in all modalities, that this is indeed occurring: we see our leg move, we hear the "woosh!" as it glides through the air, we feel it touch the object (and feel any expected level of pain), we hear the sound caused by our leg hitting the object, and we see the object itself react in accordance with our expectations. This loop is the crucial component of a convincing reality: the "reality" is virtual when the sensory data is computer generated.

2.2 Immersion

We call a computer system that supports such experience an "immersive virtual environment" (IVE). It is immersive since it immerses a representation of the person's body (the VB) in the computer generated environment. It is a virtual environment in the sense defined by Ellis [1991]: consisting of content (objects and actors), geometry and dynamics, with an egocentric frame of reference, including perception of objects in depth, and giving rise to the normal ocular, auditory, vestibular, and other sensory cues and consequences. Whether or not a system can be classified as immersive, depends crucially on the hardware, software and peripherals (displays and body sensors) of that system: we use "immersion" as a description of a technology, rather than as a psychological characterisation of what the system supplies to the human participant.

Immersion includes the extent to which the computer displays are extensive, surrounding, inclusive, vivid and matching. The displays are more extensive the more sensory systems that they accommodate. They are surrounding to the extent that information can arrive at the person's sense organs from any (virtual) direction, and the extent to which the individual can turn towards any direction and yet remain in the environment. They are inclusive to the extent that all external sensory data (from physical reality) is shut out. Their vividness is a function of the variety and richness of the sensory information they can generate [Steur, 1992].

In the context of visual displays, for example, colour displays are more vivid than monochrome, high resolution more vivid than low resolution, and displays depicting dynamically changing shadows are more vivid than those that do not. Vividness is concerned with the richness, information content, resolution and quality of the displays. Finally, as we have argued above immersion requires that there is match between the participant's proprioceptive feedback about body movements, and the information generated on the displays. The greater the degree of body mapping, the greater the extent to which the movements of the body can be accurately reproduced, and therefore the greater the potential match between proprioception and sensory data.

2.3 Presence

An IVE may lead to a sense of presence for a participant taking part in such an experience. Presence is the psychological sense of "being there" in the environment, it is an emergent property based on the immersive base given by the technology. However, any particular immersive system does not necessarily always lead to presence for all people: the factors that determine presence, given immersion, is an important area of study [Barfield and Weghorst, 1993; Heeter, 1993; Held, and Durlach, 1992; Loomis, 1992; Sheridan, 1992]. We concur with Steur [Steur, 1992] that presence is the central issue for virtual reality.
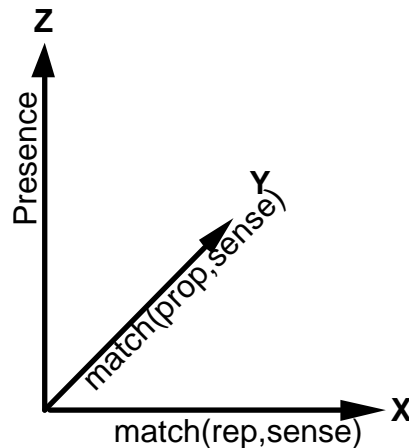
Figure 1

Presence = f(match(prop,sense), match(rep,sense)
prop = proprioception, rep = internal representation, sense = sensory data

Our view concerning the relationship between immersion and presence is shown in Figure 1. The X axis is the extent of the match between the displayed sensory data and the internal representation systems and subjective world models typically employed by the participant. Although immersion is greater the greater the richness of the displays, as discussed above, we must also take into account the extent to which the information displayed allows the particular individual to construct their own internal mental models of reality. For example, a vivid visual display system might afford some individuals a sense of "reality", but be unsuited for others in the absence of sound. Even though an excellent virtual body might exist in the VE, some individuals might reject it because it contradicts their personal self model. We have explored the relationship between presence and this match between subjectivity and displayed data in earlier experiments [Slater, Usoh and Steed, 1994].

The Y axis is the extent of the match between proprioception and sensory data, as explained above. The changes to the display must be consistent with and match through time, without lag, changes caused by the individual's motility and locomotion - whether of individual limbs or the whole body relative to the ground.

Our general hypothesis is that presence is a function of these two "matches" - that it positively increases with each of them. Note that the axes are orthogonal - a system might provide a superb degree of visual, auditory and tactile display immersion, so that most individuals have sufficient data to successfully construct their internal representations, but fail to provide a sufficient degree of match between the person's actions and the displayed results, thus breaking the link between sensory data and proprioception.

A further point about this hypothesis is that we would expect it to operate at many levels. At a very basic level, the displays should result in suitable parasympathetic responses in, for example, the ocular and vestibular systems. When an individual visually focuses on a near object the visual displays should likewise respond appropriately and immediately, and again change immediately when the focus moves to a far object. Eye tracking should be enabled. At a much higher level, when a person moves the shadow structure of the virtual body on nearby surfaces should change accordingly [Slater, Usoh and Chrysanthou, 1995]. At a similarly high level, the interactive metaphors employed in the system should match the sensory data and proprioception. This brings us back to walking: if the optical flow indicates forward movement at ground level, then the proprioceptive information should correspond to this.

A specific hypothesis of this paper is, therefore, that the degree of presence depends upon the match between proprioceptive and sensory data. The greater the match, the greater the extent to which the participant can associate with the VB as representation of self. Since the VB is perceived as being in the VE, this should give rise to a

belief (or suspension of disbelief) in the presence of self in that environment. In particular, the closer that the action required for forward locomotion corresponds to really "walking" the greater the sense of presence.

3. Locomotion

3.1 Other Methods

There is a tendency in virtual reality research to use hand gestures to do everything, from grasping objects (a natural application), to scaling the world, to navigation [Robinett and Holloway, 1992; Vaananen and Bohm, 1993]. This approach greatly overloads the hand-gesture idea - the user has to learn a complete vocabulary of gestures in order to be effective in the virtual world. Small differences between gestures can be confusing, and in any case there is no guarantee of a correspondence between the gesture and the action to be performed, and the displayed outcome.

The standard VR metaphor for locomotion is a hand gesture, with the direction of navigation determined either by gaze, or by the direction of pointing. The VPL method for navigation, as demonstrated at SIG-GRAPH 90, for example, used the DataGlove to recognise a pointing hand gesture where the direction of movement was controlled by the pointing direction.

 Song and Norman [1993] review a number of techniques, distinguishing between navigation based on eye-point movement, as opposed to object movement. Here we are interested in "naturalistic" navigation, appropriate for a walkthrough application, so we rule out navigation via manipulation of a root object in a scene hierarchy [Ware and Osborne, 1990].

Fairchild et. al. [1993] introduced a leaning metaphor for navigation, where the participant moves in the direction of body lean. The technique involves extending the apparent movement in virtual space in comparison with the real movement. In fact this is an "ice skating" metaphor, which may not be appropriate, for example, to architects taking their clients on a virtual tour.

In the context of architectural walkthrough we require participants to experience a sense of moving through the virtual building interior in a manner that maximises sensory data and proprioception. Brooks [1992] used a steerable treadmill for this purpose. However, the use of any such device as a treadmill, footpads, roller skates [Iwata, and Matsuda, 1992] or even a large area mat with sensing devices imposes constraints on the movements of participants. Moreover, there will always be an application where the virtual space to be covered is much larger than the physical space available - one of the major advantages of VR systems.

3.2 Walking

We require participants to be able to take advantage of the range available with an electromagnetic tracker, such as a Polhemus device, in order to cover small distances by moving their bodies and by really walking. Beyond the range of the sensor though, they should still carry out movements reminiscent of walking, while

staying within range. If this is possible, then proprioceptive information (associated with "walking") matches sensory data (flow in the optical array consistent with motion) to a much greater extent than motion based on hand gesture interfaces.

The new method for locomotion at ground level allows participants to move around in the space defined by the electromagnetic tracker as usual. To cover a virtual distance that is larger than the physical space afforded by the tracker, the participant *walks in place*. While carrying out this activity he or she will move forward in virtual space in the direction of gaze. It is almost walking, but no forward movement takes place in physical space. (We never have to explain to users that direction is determined by gaze, they just pick this up automatically).

A major advantage of this technique is that the hand is not used at all for navigation. The hand may be entirely reserved for the purposes for which it is used in everyday reality, that is, the manipulation of objects and activation of controls.

3.3 Implementation

The implementation of this technique is very straightforward. We have used a feed-forward neural net [Hertz, et. al., 1991] to construct a pattern recogniser that detects whether participants are "walking in place" or doing something else. The HMD tracker delivers a stream of position values (x,y,z) from which we compute first differences $(\Delta x, \Delta y, \Delta z)$ (i = 1,2,...,n). We choose sequences of n data points, and the corresponding delta-coordinates are inputs to the bottom layer of the net, so that there are 3n units at the bottom layer. There are two intermediate layers of m and m hidden units (m= m), and the top layer consists of a single unit, which outputs either 1 corresponding to "walking in place" or 0 for anything else.

We obtain training data from a person, which is used to compute the weights for the net using back-propagation. During the training phase the subject walks on the spot while immersed in almost featureless environment. He or she is asked to carry out a number of different activities, such as bending down, moving around, turning his/her head, and mixtures of these, interspersed with periods of walking on the spot. This continues for 5 to 10 minutes. An operator records binary data into the computer corresponding to whether or not the subject is walking on the spot. This data together with the corresponding sequences of delta-coordinates, are then used to train the neural net. The resulting network equations are then implemented on the virtual reality machine as part of the code of the process that deals with detection of events indicating forward movement.

After experimenting with a number of alternatives, we have found that a value of n = 20, m= 5 and m= 10 gives good results. We have never obtained 100% accuracy from any network, and this would not be expected. There are two possible kinds of error, equivalent to Type I and Type II errors of statistical testing, where the null hypothesis is taken as "the person is not walking on the spot". The net may predict that the person is walking when they are not (Type I error) or may predict that the person is not walking when they are (Type II error). The Type I error is the one that causes the most confusion to people, and is also the one that is most difficult to rectify - in the sense that once they have been involuntarily moved from where they want to be, it is almost impossible to "undo" this. Hence our efforts have concentrated on reducing this kind of error. We do not use the output of

the net directly, but only change from not moving to moving if a sequence of p 1s is observed, and from moving to not moving if a sequence of q 0s is observed (q < p). In practise we have used p = 4 and q = 2.

3.4 Results with the Neural Network

Amongst 16 people who took part in an evaluation, the mean success rate for their networks, that is the proportion of time that the net correctly predicted their activity, was 91%. The minimum and maximum rates were 85% and 96%. The mean Type I error was 10%, with minimum and maximum 6% and 15%. The corresponding figures for Type II error are 6%, 2% and 16%. Given the simplicity of the pattern recogniser we were surprised at how well the system performed in practise. We also have an arbitrarily designated "standard" network that most casual visitors to the laboratory are able to use without the necessity of a net trained for their personal style of walking.

The Polhemus Isotrak tracking device in use actually returns data to the application at a rate of about 30 Hz. The overall error is largely caused by the actual output lagging behind the real output by typically 5 samples, at the end of each sequence of 1s or 0s. It is likely that with further investigation of the neural net training method, or the employment of alternative pattern recognition techniques, results will improve.

4. Experimental Evaluation

In this Section we consider the results of two studies, a first pilot experiment and a second main experiment - each to asses the influence of the walking metaphor on ease of navigation and presence. In each case there were a number of subjects, divided equally into two groups. The first study is partially reported in [Slater, Steed and Usoh, 1993], and the second is reported here. The control groups (the "pointers") navigated the environment using a 3D mouse, initiating movement by pressing a button, with direction of movement controlled by pointing. The experimental groups (the "walkers") used the walking technique. In each case the mouse was also used for grasping objects. The task was to pick up an object, take it into a room and place it on a particular chair. The chair was placed in such a way that the subjects had to cross a chasm over another room about 20 feet below, in order to reach it.

The experiments were implemented on a DIVISION ProVision200 system. The ProVision system includes a DIVISION 3D mouse, and a Virtual Research Flight Helmet as the head mounted display. Polhemus sensors are used for position tracking of the head and the mouse. Scene rendering is performed using an Intel i860 microprocessor (one per eye) to create an RGB RS-170 video signal which is fed to an internal NTSC video encoder and then to the displays of the Flight Helmet. These displays (for the left and right eye) are colour LCDs with a 360 × 240 resolution and the HMD provides a horizontal field of view of about 75 degrees. The frame update rate achieved during the experiments was about 15 frames per second.

All subjects saw a virtual body as self representation. They would see a representation of their right hand, and their thumb and first finger activation of the 3D pointer buttons would be reflected in movements of their corresponding virtual finger and thumb. The hand was attached to an arm, that could be bent and twisted in response to similar movements of the real arm and wrist. The arm was connected to an entire but simple block-like body representation, complete with legs and left arm. Forward movement was accompanied by walking motions of the virtual legs. If the subjects turned their real head around by more than 60 degrees, then the virtual body would be reoriented accordingly. So for example, if they turned their real body around and then looked down at their virtual feet, their orientation would line up with their real body. However, turning only the head around by more than 60 degrees and looking down (an infrequent occurrence), would result in the real body being out of alignment with the virtual body.

4.1 Navigation

With respect to ease of navigating the environment, subjects in both experiments marginally preferred to use the pointing technique. This result was not surprising: as Brooks noted with the real treadmill, certainly more energy is required to use the whole body in a walking activity, compared to pressing a mouse button, or making a hand gesture (or driving a car, with respect to the similar comparison in everyday reality). Moreover, the networks did not work with 100% accuracy, in contrast to the accuracy of the pointing method.

In the post-experiment questionnaire three questions were asked of all subjects, covering three aspects of navigation: general movement, that is how simple or complicated it was to move around; placement, that is the ease of getting from one place to another, and how "natural" the movement was. The questions are shown in Table 1, with results given for both experiments (the results should not be combined since there were some differences between the two experiments). The differences between the answers given by the "pointers" and "walkers" are not statistically significant. However, Figure 2 shows scattergrams, for those in the experimental group of the answers to the three questions against the Type I error for the pilot study only (such data was not available from the main study). The sample size involved is too small to carry out meaningful significance tests, but the graphs indicate that a decrease in Type I error generally leads to an improvement in ease of navigation. This suggests that a better pattern recognition technique could result in a superior performance for this method of navigation, compared to the pointing method. In other words, it is worth-while improving the pattern recognition technique, for decrease in error is likely to result in a substantial improvement in subjective evaluation. (With the pointing technique there is no similar improvement that can be made).

Table 1

Questions Relating to Ease of Navigation

General Movement		Getting from place to place		Natural/unnatural
Did you find it relatively "simple" or relatively "complicated" to move through the computer generated world?		How difficult or straightforward was it for you to get from place to place?		The act of moving from place to place in the computer generated world can seem to be relatively "natural" or relatively "unnatural". Please rate your experience of this.
To move through the world was...		To get from place to place was ...		The act of moving from place to place seemed to me to be performed ...
1. very complicated		1. very difficult		1. very unnaturally
...		...		...
7. very simple		7. very straightforward		7. very naturally
PILOT STUDY				
Mean Response		Mean Response		Mean Response
Control Group: 5.0, n = 6		Control Group: 4.9, n = 6		Control Group: 3.4, n = 6
Exp. Group: 5.1, n = 8		Exp. Group: 5.5, n = 8		Exp. Group: 3.9, n = 8
MAIN STUDY				
Control Group: 5.5, n = 8		Control Group: 5.7, n = 6		Control Group: 4.2, n = 6
Exp. Group: 4.9, n = 8		Exp. Group: 4.7, n = 8		Exp. Group: 4.2, n = 8

Movement Natural

Type I error

(a)

General Movement

Type I error

(b)

Getting from Place to Place
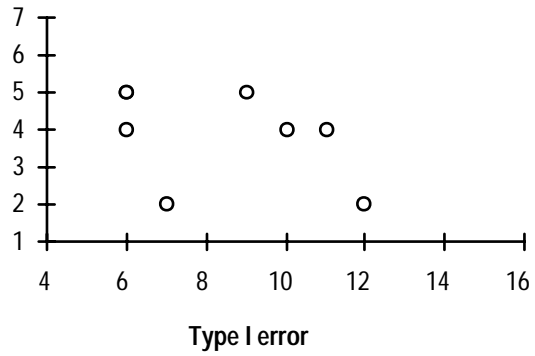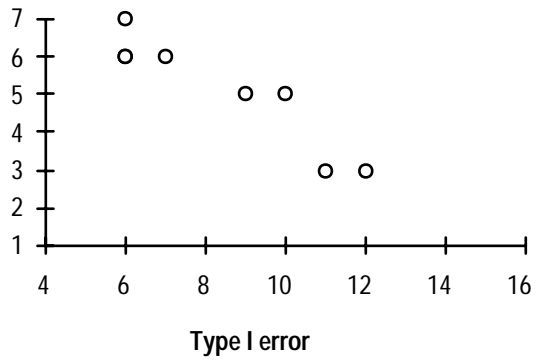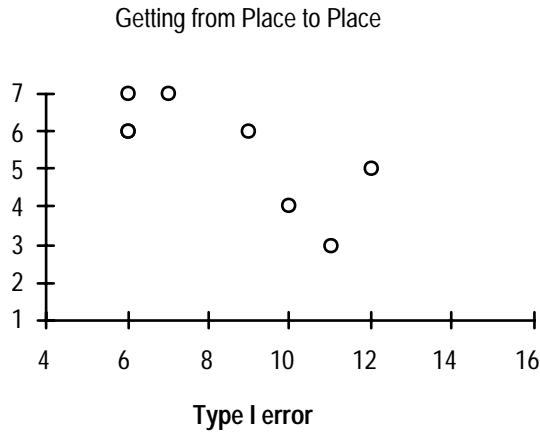

(c)

Figure 2

Evaluation of Navigation by Type I Error

4.2 Presence

It is the sense of presence with which we are mainly concerned. Here we discuss the results of the main experiment that compared the two different techniques for navigation with respect to the effect on reported sense of presence.

There were 16 subjects, divided into two groups of eight. These were selected by asking for volunteers on the QMW campus, excluding people who had experienced our virtual reality system before, or who knew of the purposes of our research. The control groups (the "pointers") moved through the environment using the DIVISION 3D mouse, by pressing a button, with direction of movement controlled by pointing. The experimental groups (the "walkers") used the walking technique. All subjects used the same ("standard") network based on the "walking in place" behaviour of one individual. Both walkers and pointers used the mouse for grasping objects. Intersecting the virtual hand with an object and pulling the first finger (trigger) button, resulted in the object being attached to the hand. The object would fall when the trigger button was released.

The task in the experiment was to pick up an object located in a corridor, take it into a room and place it on a particular chair. The chair was placed in such a way that the subjects had to cross a chasm over another room about 20 feet below, in order to reach it. The subjects could get to the chair either by going out of their way to walk around a wide ledge around the edges of the room, or by directly moving across the chasm. This was a simple virtual version of the famous visual cliff experiment [Gibson and Walk, 1960].

Subjective presence was assessed in three ways: the sense of "being there" in the VE, the extent to which there were times that the virtual world seemed more the presenting reality than the real world, and the sense of visiting somewhere rather than seeing images of something. Each was rated by subjects on an ordinal 7 point scale, where 7 was the highest score, using a questionnaire given immediately after the experiment. These three scores were combined into one by counting the total number of 6 or 7 responses from the three questions. Hence, the result was a value between 0 and 3.

Other questions relevant to the analysis concerned the degree of nausea experienced in the VR, and the extent of association with the VB: ("To what extent did you associate with the computer generated limbs and body as being 'your body' while in the virtual reality?"). They were also asked to rate the degree of vertigo, if any, induced by the virtual precipice, and also to compare their reaction to this in relation to how they would have reacted to a similar situation in real life. ("To what extent was your reaction when looking down over the drop in the virtual reality the same as it would have been in a similar situation in real life?").

All subjects were watched by an observer, who, in particular, recorded whether or not they moved to the chair by walking around the ledge at the side of the room, or by walking directly across the precipice. In the event, only four subjects out of the sixteen (two from each group) walked across the precipice.

The main conclusion from the statistical analysis was that for the "walkers", the greater their association with the VB the higher the presence score, whereas for the "pointers" there was no correlation between VB association and the presence score. In other words, participants who identified strongly with the virtual body had a greater degree of reported presence if they were in the "walking" group than if they were in the "pointing" group. Association with the VB is important. This certainly belongs to the X axis of Figure 1: indicating that it is not simply a question of whether a VB is provided by the system and how well it functions, but also the individual's personal evaluation of this VB, the degree of "match" to their internal world models. It also belongs to the Y axis, as is discussed in Section 7 below.

There were two other statistically significant factors. First, path taken to the chair: a path directly over the precipice was associated with lower presence. This is as would be expected, and is useful in corroborating the veracity of the presence score. Second, degree of nausea: a higher level of reported nausea was associated with a higher degree of presence. This same result has been found in each of our studies. We speculate that the vection in VR is a cause of both simulator sickness and an influence on presence [McCauley & Sharkey, 1993]. Finding nausea and presence associated would therefore not be surprising. There is the further point that presence is concerned with the effect of the environment on the individual. An increased sense of presence is likely to be correlated with the human brain paying more attention to the detailed operation of the environment, and therefore to the dis-

crepancy between the visual and vestibular systems. However, this may be a temporary phenomenon that is overcome with greater exposure. This is speculation, and would need to be examined by an independent study.

5. Statistical Analysis for Presence

The dependent variable (p) was taken as the number of 6 or 7 answers to the three questions as stated above. The independent variable was the group (experimental or control). The explanatory variables were VB, degree of association with the Virtual Body; S the reported nausea, and P for path (= 1 for a path around the sides of the room, and 2 for a direct path across the precipice).

This situation may be treated by logistic regression [Cox, 1970], where the dependent variable is binomially distributed, with expected value related by the logistic function to a linear predictor where N (=16) is the number of observations and where n (=3) is the number of binomial trials per observation.

Table 2 shows the results. The overall model is significant. For a good fit, the overall deviance should be small, so that a value of less than the tabulated value is significant. No term can be deleted from the model without significantly increasing the deviance (at the 5% level).

The analysis relies on the assumption that the dependent variable is binomially distributed. This assumption is made as a heuristic, but cannot be justified in an obvious way. The presence-related questions were each separated by at least three others in the questionnaire, and for any respondent, not knowing the purposes of the study, and not aware of the concept of presence, it would be reasonable to assume that their answers did not directly influence one another, and therefore that the "trials" were independent.

Table 2

Logistic Regression Equations

= fitted values for the presence scale

VB = VB association, S = Nausea, P = Path

Group	Model	When P=2 (path directly over precipice)
Walkers	= -16.9 + 2.6*VB + 1.3*S	-2.7
Pointers	= -3.1 + *0.1*VB + 1.3*S	-2.7

Non-significant coefficients are shown in *italics*.

Overall Deviance =11.424, d.f. = 10
χ2 at 5% on 10 d.f. = 18.307

Deletion of Model Term	Change in Deviance	Change in d.f.	cat 5% level
S	6.624	1	3.841
P	3.867	1	3.841
Group.VB	10.922	2	5.991

An alternative analysis was carried out, where the three presence scores were combined into a single scale using principal components analysis [Kendall, 1975]. The first principal component is the linear combination of the original variables that maximises the total variance. The second is orthogonal to the first and maximises the total residual variance. The first two principal components accounted for 96% of the total variation in the original three variables (the first for 67% and the second for 29%). The single presence score was taken as the norm of the vector given by the first two principal components.

Table 3

Regression Equations

= fitted values for the presence scale based on principal components
(Coefficients are given to 1 d.p.)

VB = VB association, S = Nausea, C = Vertigo comparison

Group	Model	When C=2 "same as real life"
Walkers	= -4.5 + 1.7*VB + 1.2*S	+ 2.5
Pointers	= 3.4 + *0.3*VB + 1.2*S	+ 2.5

Non-significant coefficients are shown in *italics*.

A regression analysis using this new presence score resulted in a model qualitatively similar to that described above. Here though, instead of P (path) being significant, the variable representing the comparison between vertigo experienced in the virtual world with what might have been experienced in the real world, was significant instead. A higher degree of presence was associated with the comparison resulting in a "same as real life". The overall regression was significant at 5% with a multiple squared correlation coefficient of 0.81. This is summarised in Table 3.

6. Steps and Ladders

6.1 Walking on Steps and Ladders

In the previous sections we have made a case, together with supporting experimental evidence, that the walking in place technique tends to increase subjective presence, in comparison with the pointing technique based on a simple hand gesture, provided that there is an association with the VB.

The same idea can be applied to the problem of navigating steps and ladders. One alternative is to use the familiar pointing technique and to fly. While in some applications there may be a place for such magical activity, the very fact that mundane objects such as steps and ladders are in the environment would indicate that a more mundane method of locomotion be employed. The walking in place technique carries over in a straightforward manner to this problem.

When the collision detection process in the virtual reality system detects a collision with the bottom step of a staircase, continued walking will move the participant up the steps. Walking down the steps is achieved by turning around, and continuing to walk. If at any moment the participant's virtual legs move off the steps (should this be possible in the application) then they would "fall" to the ground immediately below. Since walking backwards down steps is something usually avoided, we do not provide any special means for doing this. However, it would be easy to support backwards walking and walking backwards down steps by taking into account the position of the hand in relation to body line: a hand behind the body would result in backwards walking.

Ladders are slightly different; once the person has ascended part of the ladder, they might decide to descend at any moment. In the case of steps the participant would naturally turn around to descend. Obviously this does not make sense for ladders. Also, in climbing ladders it is usual for the hands to be used. Therefore, in order to indicate ascent or descent of the ladder, hand position is taken into account. While carrying out the walking in place behaviour on a ladder, if the hand is above the head then the participant will ascend the ladder, and descend when below the head. Once again it is a whole body gesture, rather than simply use of the hand that is required in order to achieve the required result in an intuitive manner. If at any time the virtual legs come off the rungs of the ladder, then the climber will "fall" to the ground below.

6.2 Evaluation for Usability

We have only carried out a simple study to test for usability. A scenario was constructed consisting of steps leading up to the second storey of a house. The steps led in through a doorway, which entered into a room consisting of a few everyday items such as a couch, TV, and so on. There was a window, and a ladder down to the ground outside propped up against the wall just below the window. There was a bucket on the ground outside, at the foot of the ladder. Examples are shown in the colour Plates.

The task was to walk up the steps, enter into the room, climb onto the ladder and down to the ground, pick up the bucket, take it back up into the room, down the stairs, and back outside. The designer of this scene was taken as the "expert" - and completed the scenario in 3 minutes, including one fall from the ladder. Five other people, all of whom had used the VR system before, were invited to try out the scenario. One person also completed the task in 3 minutes, without any falls. Another took 4 minutes, also without any falls. The third required 5 minutes with 2 falls from the ladder. The remaining two each took 8 minutes, with 1 and 2 falls from the ladder respectively. The results of this simple experiment were encouraging enough for us to consider devising specific pattern recognisers for these types of activities.

7. Conclusions

The rudimentary model for presence in virtual environments illustrated in Figure 1 forms a context in which the walking in place technique for locomotion at ground level can be considered. We argue that the walking technique is a shift along the Y axis of Figure 1, compared to the pointing technique, and therefore other things being equal, should result in a greater sense of presence. However, we found that this is modified by the degree of association of the individual with the virtual body. In fact this factor spans both X and Y axes: lack of association may be due to lag between real and displayed virtual movements (Y axis), or immobility of the virtual left hands and feet (Y axis), or to the rather simple visual body model (X axis). In any case, the VB association is significantly positively correlated with a subjective presence for the walkers, but not for the pointers, which is certainly consistent with the proposed model.

In earlier work [Slater and Usoh, 1994] we used the term "body centred interaction" for techniques that try to match proprioception and sensory data. The walking in place method is a clear example of this. When the method works well it feels like walking and the corresponding flow in the optical array matches both head movements and the movements of the feet. Also, the technique is very easy to understand for there is little to learn as such, and therefore this is less of a *metaphor* than other techniques. In this case we walk by "almost walking", rather than doing some other activity that is completely different to walking and then having to make the mental association between the cause and effect. The empirical evidence does not support the notion that people prefer this for navigation compared to pointing, but it does suggest that improved performance of the neural net based pattern recogniser may lead to such a preference.

We have described the technique applied to climbing or descending steps and ladders. This may be useful in circumstances where the interaction style should be relatively mundane, rather than requiring magical effects such as "flying". Training for fire fighting, the application that inspired the extension to steps and ladders, clearly falls into this category.

Body Movement in the Garden

Summary

This chapter describes an experiment to assess the influence of body movements on presence in a virtual environment. Twenty subjects were required to walk through a virtual field of trees and count the number of trees with diseased leaves. A 2×2 between-subjects design was used to assess the influence of two factors on presence. One factor was tree height variation, and a second factor was the complexity of the task. The field with greater variation in tree height required subjects to bend down and look up more than those in the lower variation tree height field. Those with the higher complexity task were told to remember the distribution of diseased trees in the field, as well as to count them. The results showed a significant positive association between reported presence and the amount of body movement, in particular head yaw, and the extent to which subjects bent down and stood up. There was also a strong interaction effect between the task complexity and gender, with females in the group with the more complex task reporting a much lower sense of presence than in the simpler task.

1. Introduction

When information about an environment is presented to an individual, that individual may have a sense of being *present* in that environment to a greater or lesser extent. In particular, when someone receives sensory data from one environment (say, kinesthetic and tactile information from the real world) and different, perhaps contradictory, data from a competing environment (say, visual and auditory data from a computer generated virtual environment) they may be more or less present in each environment. Presence refers to the sense of 'being in' an environment, and only

really makes sense when speaking about the degree of presence in one environment relative to another (Slater, Usoh, Steed, 1994). This is because a conscious individual receiving sensory stimuli from only one environment is, by definition, present there. When competing stimuli are simultaneously received from multiple environments there is the issue of which, if any, comes to dominate and why. The main hypothesis of the experiment carried out in this paper is that the environment relative to which major body movements are made has a higher probability of being the dominant presence environment, other things being equal.

Presence may be subjective or behavioural. Subjective presence refers to what an individual will express in response to questions about 'being there'. Behavioural presence refers to observable responses to stimuli. Although related in practice, there is no necessary logical connection between these two - we think of subjective presence as being a verbal and necessarily conscious articulation of a state of mind, and behavioural presence as being automatic, unplanned non-conscious bodily responses. Both types are important: subjective presence is essentially an evaluation of an experience, whereas behavioural presence is concerned with responses to events in the environment in question - clearly important in applications such as training or psychotherapy (Hodges et. al., 1996). Likewise, presence may be measured by subjective means such as questionnaires, or through observation of behaviour.

Presence research has focused on definition and ideas for measurement (Heeter, 1992; Held and Durlach, 1992; Loomis, 1992; Sheridan, 1992, 1996; Steur, 1992; Ellis, 1996; Slater and Wilbur, 1997; Zeltzer, 1992) and there have been several empirical studies of contributing factors (Barfield and Weghorst, 1993; Barfield et. al., 1995; Hendrix and Barfield, 1996a, 1996b; Slater et. al., 1995; Welsh et. al., 1996). Some of the factors studied have included the effect of visual display update rate, characteristics of the visual display system, the influence of spatialised sound, head-tracking, and interaction.

This paper describes an experiment to examine the influence of two factors on subjective presence in virtual environments - the extent of body movement, and also complexity of a task undertaken in the environment. The major interest is on body movement which has important practical consequences for the design of interactive paradigms in VEs.

This study is motivated by two considerations. The first is anecdotal - we have observed hundreds of subjects in head-mounted display based 'virtual reality' over the past six years. It is very frequently the case that when a person first dons the head-mounted display (HMD), that they treat it as a 'computer screen' and just stand rigidly looking ahead. When they are told to move - turn their head, bend down, reach up, look under, they frequently have an observable 'aha!' type experience indicating a transition from low to high presence. It is this effect of body movements on presence that is explored via the experiment reported in this paper.

The second motivation is a practical one. The goal is to construct interactive techniques that exploit the idea of whole body gestures, in order to maximise presence. An assumption underlying previous work on 'body centred interaction' (Slater and Usoh, 1994) has assumed that whole body movement semantically appropriate to the task will enhance presence. This has been experimentally tested in the context of ground based locomotion

(Slater, Usoh, Steed, 1995), but not in a more general setting. This is the major purpose of the work described here.

2. Subjective Presence

The vast majority of studies measure presence through questionnaire, and are thus eliciting subjective presence. Witmer and Singer (1998) have developed an extensive Presence Questionnaire. However, their approach mixes what we have called 'immersion' (the objective factors such as field of view, display resolution, or degree of interactivity possible) and the psychological and behavioural response to these factors that we term as 'presence' (Slater and Wilbur, 1997). For example, their first question asks 'How much were you able to control events?' We see this as eliciting the subject's view of one of the aspects of immersion, rather than being directly concerned with presence. Therefore for this study we preferred to use a questionnaire and methodology that we have used for several previous experiments (for example, Slater, Usoh, Steed, 1995).

There are six questions each on a 1 to 7 scale, where the higher score always means higher reported presence. A conservative measure of subjective presence is then constructed as the number of high responses (scores of 6 or 7) in the answers to the six questions. Under the null hypothesis that scores are attributed randomly and independently, this results in a binomially distributed count (number of high responses out of 6) as the response variable, and logistic regression (McCullagh and Nelder, 1983, Chapter 4) can then be used to analyse the responses. This method is preferred on statistical grounds because it does not involve treating the ordinal response data in any way as if it were interval data, and is appropriately conservative in measuring subjective phenomena.

The particular questions used in the current study, scattered throughout a larger questionnaire, were as follows:

1. Please rate *your sense of being in the field amongst the plants*, on the following scale from 1 to 7, where 7 represents your *normal experience of being in a place*.
I had a sense of "being there" in the field:
1. Not at all ... 7. Very much.

2. To what extent were there times during the experience when the virtual field of plants became the "reality" for you, and you almost forgot about the "real world" of the laboratory in which the whole experience was really taking place?
There were times during the experience when the virtual field became more real for me compared to the "real world"...
1. At no time ... 7. Almost all the time.

3. When you think back about your experience, do you think of the virtual field more as images that you saw, or more as somewhere that you visited? Please answer on the following 1 to 7 scale:
The virtual field seems to me to be more like...
1. images that I saw ...7. somewhere that I visited.

4. During the time of the experience, which was strongest on the whole, your sense of being in the virtual field, or of being in the real world of the laboratory?
I had a stronger sense of being in...
1. the real world of the laboratory ... 7. the virtual reality of the field of plants.

5. Consider your memory of being in the virtual field. How similar in terms of the structure of the memory is this to the structure of the memory of other places you have been today? By 'structure of the memory' consider things like the extent to which you have a visual memory of the field, whether that memory is in colour, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic in your imagination, and other such structural elements.
I think of the virtual field as a place in a way similar to other places that I've been today....
1. not at all ...7. very much so.

6. During the time of the experience, did you often think to yourself that you were actually just standing in an office wearing a helmet or did the virtual field of plants overwhelm you?
During the experience I often thought that I was really standing in the lab wearing a helmet....
1. most of the time I realised I was in the lab ... 7. never because the virtual field overwhelmed me.

3. Method

3.1 Factorial Design

The overall purpose of the experiment was to assess the extent to which body movement, in particular bending down, and turning the head around and up and down, influences presence. A scenario was devised that naturally would induce some subjects to use bending and head movements more than others.

The scenario consisted of a field of unusual plants or trees with large leaves, distributed at random through the field (Figure 1). Half the subjects were put into a field where the heights of the trees varied considerably, some being much below head height and some very much taller. The other subjects, were put into a field where the tree heights were all above normal standing eye level. Healthy plants had green leaves. Diseased plants could be distinguished from healthy ones because the *underneath* of their leaves were discoloured (brown). Moreover, for the trees in the high variation field the leaves were folded inwards in such a way that it would only be possible to see their underneath by looking upwards while underneath the tree. For the low variation field the leaves were arranged in such a way that it was possible to see their underneath by looking approximately at eye height in a standing position.

Figure 1 about here

All subjects were asked to move through the field in any direction they preferred, and to count the number of diseased plants. A more complex task was also given to some subjects, not only to count the number of diseased plants but *also to remember* where they were in order to later draw a map showing their distribution throughout the field. The purpose was to examine whether the more complex task would affect presence.

There were 20 subjects in total, and a between subjects factorial design was used with five subjects in each condition. The subjects were recruited by the Department of Psychology and paid £5 (about $9 US) each for completion of the full experiment and all questionnaires. Most of the subjects were students (3 undergraduate, 8 Masters, 4 PhD), and there were 3 Research Assistants, 1 member of the administration and 1 journalist. There

were 13 male subjects. No subject had any involvement in the research or any knowledge of the purpose of the experiment.

There were 150 trees in each scene, randomly distributed in a garden of dimension 90m×75m. Each tree was 2.4m across, and had 16 leaves. There were three classes of tree in equal proportions (50 each), one healthy, one with 1 bad leaf, one with 4 bad leaves. For the low variation field the distribution of heights was 1.7m ± 0.1m, and 2.35m ± 1.9m for the high variation field.

3.2 Procedures

When the subjects donned the HMD they were placed in a virtual environment that was a rendition of the *same laboratory in which they were actually standing*. The experimenter continued to refer to what they were experiencing as "being in the lab" where they carried out some initial training tasks.

After this short training session, the subjects were asked to look around "the lab", and instructed to turn their head, bend down, stand up, so that they realised that these actions were possible. Then they were asked to turn around 180 degrees, and locate the door to the lab. They were told that when the door opened they should go through it, and they would enter the field of plants. Hence the field was located beyond the door, and from any position in the field it was possible to see the door back to the lab. The subjects then went into the field and carried out their task. This continued for about 3 minutes. They were told beforehand that they were to begin to make their way back to the lab, though still continuing with their task, once the sky became brighter (the sky started off as black, but after 3 minutes it became light blue). From earlier pilot experiments it had been found that about 3 minutes was the right length of time after which many subjects started to become visibly bored by the task.

During the time that they were in the virtual field, the experimenter said nothing. On returning back to the 'lab', the experimenter said "Welcome back! Well done!" and continued to talk as if they were back in "the lab". After another short set of tasks the subjects were asked to look around the lab once again, and then the HMD was removed, and again they were asked to look around the lab. After this the questionnaires were administered.

3.3 Explanatory Variables

- Information was collected on many explanatory variables, the most relevant for this paper being:
- background information such as gender and occupation.
- pitch in degrees/sec. - the summation (after smoothing for noise) of all vertical (i.e., pitch) angles through which the head moved. (i.e., project the head orientation vector onto a vertical plane, and measure the angle between two successive head orientations. pitch is the sum of all such successive angles divided by time).
- yaw in degrees/sec. - a similar measure for yaw angle - the sum of horizontal angles through which the head turned divided by time.
- roll in degrees/sec. - a similar measure for roll angle.
- mean and standard deviation of hand height above ground level (m).

3.4 Materials

The scenarios were implemented on a Silicon Graphics Onyx with twin 196 MHz R10000, Infinite Reality Graphics and 64M main memory. The software used was Division's dVS and dVISE 3.1.2. The tracking system has two Polhemus Fastraks, one for the HMD and another for a 5 button 3D mouse. The helmet was a Virtual Research VR4 which has a resolution of 742×230 pixels for each eye, 170,660 colour elements and a field-of-view 67 degrees diagonal at 85% overlap.

The total scene consisted of 32,576 triangles (almost all of these accounted for by the 150 trees) which ran at a frame rate of no less than 10Hz in stereo. The display lag was approximately 100ms.

Subjects moved through the environment in gaze direction at constant velocity by pressing a thumb button on the 3D mouse. Subjects had a simple inverse kinematic virtual body. Most of the time they remained unaware of their virtual arm and hand because of the relatively limited field of view.

4. Results

4.1 Measuring Body Movement

The fundamental question concerns the relationship between body movement and presence. Body movement in the conditions of this experiment has two components: the degree to which there was variation in the whole body height (the extent of bending down and standing up), and also the degree of head rotation. As it turned out, the first factor 'tree height' was the main source of variation for the first type of movement, and the second factor 'task' was the major source of variation for the second.

For the low variation tree field, although there was no bending down, there was considerable head rotation, for most of the subjects did not realise that when approaching the trees from certain directions that it was possible to see the underneath of the leaves without having to rotate the head upwards. The average head movement (rotation in pitch, yaw and roll) was not significantly different between the two tree groups - the major difference being the amount of overall body movement.

A measure of how much the body crouched down and stood up (independently of head rotation) was obtained by measuring hand height. The hand was not required to do anything other than hold the 3D mouse and press a button for locomotion. Therefore the changes in height reflected changes in overall body extension. The variable used for this was the ratio of mean to standard deviation in hand height. This would be smaller for subjects who tended to bend down and stand up more than those who remained standing (or sitting) throughout. The mean 'hand ratio' was 13.7 ± 12.2 for those in the low variation tree group, and 3.8 ± 2.0 for those in the high variation group. This numerical variable was therefore used *instead* of the 'tree' factor itself in the analysis, in order to avoid the problems caused, for example, by (the two) subjects who sat on the floor throughout.

The secondary question concerned the relationship between task complexity and presence. There were two levels of the factor 'task', level 1 corresponding to the instruction just to count the diseased trees, and level 2 to count and remember to later sketch the distribution of diseased trees. This factor was confounded with head rotation

(pitch, yaw and roll rotation). There are significant differences between the mean head rotations between task levels 1 and 2, with the more complex task leading to higher rotation, especially yaw, as shown in Table 1.

Table 1 about here.

4.2 Logistic Regression

A binomial logistic regression analysis was used for the 'presence' response, which is the count of high scores out of six questions. Logistic regression is a standard technique for the analysis of binomial data, and involves a logistic function transformation in order to ensure that the fitted values are within the range of allowable values (between 0 and the maximum possible count - being 6 in this case).

Suppose, for example, that there were two factors, A and B with h_A levels of factor A and and h_B levels of factor B. In the (i,j)th cell there are n_{ij} responses $(y_{ijk}, k = 1,...n_{ij})$, and suppose that associated with each are two explanatory variables x_{ijk} and z_{ijk} ($i = 1,...,h_A; j = 1,...,h_B; k = 1,...,n_{ij}$). Then the linear predictor is of the form:

$$\eta_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + b_{ij} x_{ijk} + c_{ij}z_{ijk}$$

....(1)

$$i = 1,...,h_A;$$
$$j = 1,...,h_B;$$
$$k = 1,...,n_{ij}$$

where α_i is the main effect for factor A, β_j the main effect for factor B, and γ_{ij} is the interaction effect between A and B. The model also allows the regression slopes (b_{ij} and c_{ij}) to be different across the factor levels. Solution to the least squares equations requires constraints on the parameters, achieved by setting the first level of each coefficient to zero: $\alpha_1 = \beta_1 = \gamma_{i1} = \gamma_{1j} = b_{1j} = b_{i1} = c_{1j} = c_{i1} = 0$. This is the standard approach for such generalised linear models (McCullagh and Nelder, 1983).

The logistic regression links the expected value of the presence count $E(y_{ijk})$ to the linear predictor as:

$$E(y_{ijk}) = \frac{n}{1 + \exp(-\eta_{ijk})}$$

where n (=6) is the number of binomial trials per observation (the six presence questions). Maximum likelihood estimation is used to obtain estimates of the coefficients, which is equivalent to iteratively reweighted least squares on the transformed response variate η. The deviance (minus twice the log-likelihood ratio of two models) may be used as a goodness of fit significance test, comparing the null model (all coefficients are zero) with any given model. The change in deviance for adding or deleting groups of variables may also be used to test for their significance, and in the following sections all significance tests are at the 5% level. The (change in) deviance has an approximate χ^2 distribution with degrees of freedom dependent on the number of parameters (added or deleted). A good overall fit should result in a low deviance (judged against the corresponding chi-squared value).

4.3 The Fitted Model

A number of different models were compared by starting from the baseline model that included only task and the hand height mean/standard deviation ratio, and then adding and deleting terms. A very good overall fit was obtained with a model that additionally included the interaction between task and gender, and the head-rotational yaw variable (Table 2). Note that the range of the yaw variable is from 6.6 to 17.9 deg./sec., and the range of the hand height ratio is from 0.34 to 33.7.

Table 2 about here.

The coefficients shown for main and interaction effects are the *changes* to the constant induced by introduction of the corresponding term. The overall fitted model, expressed as a linear model for the predictor (η) is shown in Table 3.

Table 3 about here.

4.4 Analysis

Task by itself is not significant. However, there is a very significant interaction effect between gender and task. For males the mean presence count is 2.2 ± 2.1 for the simpler task (1), and 3.4 ± 1.6 for the more complex task (2). For females the means are 5.2 ± 0.5 (task 1) and 0.7 ± 0.6 (task 2).

The results do show a significant relationship between presence and the body movement variables. It is positively associated with yaw, and negatively associated with the amount of vertical variation (as measured by the hand height mean to standard deviation ratio). In other words those who had a lower mean hand height and greater variation reported higher presence than those with a higher mean and lower variation, other things being equal.

The result for yaw can be queried on the grounds of the confounding between yaw and the task factor. However, the inclusion of an additional interaction term between yaw and task is just above the 5% significance borderline, and shows that for those in the more complex task (task 2) group there is a positive association between yaw and the presence count. Inspection of these results yielded two outliers, and when removed the inclusion of a yaw/task interaction term was not significant. It is safe to conclude that the impact of head turns holds independently of the task effect. Finally, either roll or pitch head movements could be included (significantly) instead of yaw, but not in addition to yaw; i.e., only one of these three variables could be included and yaw was the most significant by far.

5. Conclusions

In this paper we have considered an issue of theoretical and practical importance. The theoretical issue is really that of the existence of the phenomenon of 'presence' at all. Some researchers in virtual reality have taken presence as a central issue - essentially as a guide as to what constitutes a 'good' virtual reality system (within a par-

ticular application context). A good system is, in this view, one that delivers greater presence, not only because of the evaluative aspects but also because higher presence should lead to behaviour in the VE being similar to what it would be in everyday reality in comparable circumstances.

The practical importance of the results of this experiment is that since there does seem to be a relationship between body movement and presence, it is a reasonable goal to design interactive paradigms that are based on semantically appropriate whole body gestures. These will not only seem more 'natural', but may also increase presence. We further believe that the increase in presence in itself will engender more body movement, which in turn will generate higher presence, and so on.

The idea of a transition from the real lab, to a virtual lab, to the experimental scenario, back to the virtual lab, and then to the real lab, may prove a useful means for easing subjects into the virtual environment. The virtual lab may be thought of as a sort of 'presence ante-room'. It could be used to prepare the subjects for the experiment, and then as a 'place' in which to measure presence when they return. It provides experimenters with a way of continuing to talk to the subjects even after they have entered the VE. It provides the opportunity for pre-and post-experimental measurements to be taken while in the VE. There is a lot more data that was collected that is relevant to this issue, which will be presented in further reports.

The conclusions of this paper are that the reported presence of a participant in an immersive VE is likely to be positively associated with the amount of whole body movement (such as crouching down and standing up), and head movements (looking around and looking up and down) appropriate to the context offered by the VE. The experiment also considered the impact of a certain type of task complexity on presence, but the results in this case were inconclusive because of a confounding of the task with head-rotation, and a strong interaction effect between task and gender. This paper has concentrated only on subjectively reported presence; further work should examine whether the results extend to behavioural presence.

TABLE 1
Means and Standard Deviations of the Head-movement Variables by Factor
(Differences between each pair of Task means are significant on a t-test at 5% on 18 df).

Variable	Tree 1 (low)	Tree 2 (high)	Task 1 (simple)	Task 2 (complex)
pitch deg./sec.	3.4 ± 1.4	3.3 ± 1.4	3.7 ± 0.9	5.1 ± 1.1
yaw deg./sec.	12.3 ± 3.7	10.1 ± 3.5	9.4 ± 3.3	13.0 ± 3.1
roll deg./sec.	13.5 ± 4.1	11.2 ± 3.7	2.6 ± 0.7	4.0 ± 1.7

TABLE 2

Logistic Regression for Subjective Presence

Overall deviance = 21.600, df =14, χ^2 = 23.685 at 5%

Δdev is the change in deviance caused by deletion of the corresponding term, and has a Chi-Squared distribution with 1df in each case.

χ^2 = 3.841 on 1df at 5%.

The Symbol column shows the corresponding term in the model from Eq.(1).

Parameter	Symbol	Estimate	S.E.	Δdev
Constant	μ	-2.66	0.82	
Main Effects:				
Task	α_2	0.47	0.54	
Gender female	β_2	3.19	0.77	
Interaction Effects:				
Task and Gender	γ_{22}	-6.09	1.18	37.5
Slopes				
yaw	b	0.24	0.074	12.1
hand height ratio	c	-0.056	0.026	4.9

TABLE 3
Overall Logistic Regression Model
Linear Predictor η = Const. + 0.24 yaw - 0.06 height
where the constant is given below:

Male		Female	
Task 1	**Task 2**	**Task 1**	**Task 2**
-2.66	-2.20	0.52	-5.10

Counting Presence in the 3D Chess World

Summary

This chapter describes a new measure for presence in immersive virtual environments (VEs) based on data that can be obtained unobtrusively during the course of a VE experience. At different times during the course of an experience a participant will perceive the VE generated sensory data as the foreground, with sensory data from the real outside laboratory as the ground, and occasionally switch between these two interpretations of their total sensory input. The number of transitions from virtual to real is counted, and using some simplifying assumptions, a probabilistic Markov Chain model can be constructed to model these transitions. This can be used to estimate the equilibrium probability of being 'present' in the VE. This technique was applied in the context of an experiment to assess the relationship between presence and body movement in an immersive VE. The movement was that required by subjects to reach out and touch successive pieces in a Tri-Dimensional chess board. The experiment included 20 subjects, 10 of whom had to reach out to touch these chess pieces ('active group'), and the other 10 controls only had to click a hand-held mouse button. The results showed that amongst the active group there was a significant positive association between body movement and presence. The result lends support to interaction paradigms that are based on maximising the match between sensory data and proprioception.

1. Introduction

One way to evaluate whether a virtual environment (VE) system is 'working' is the degree to which it generates a sense of 'presence' in its participants. Presence occurs when sensory data from the VE perceptually dominates the all-pervasive (and often conflicting) data from the real world environment in which the whole experience is actually taking place. This gives a person a sense of 'being' in the place depicted by the VE. The behavioural concomitant of virtual presence could be argued to be crucial to certain applications - for example, the use of VEs in psychotherapy would not be feasible unless clients developed a sufficient state of 'being there' (on the bridge, in the aircraft, in the elevator) as to elicit the types of response that would occur in similar real experiences (Rothbaum et. al., 1995; Hodges, et. al., 1996). It can be argued that presence is crucial in any application that might be considered as a 'virtual rehearsal' for something in real life.

This paper introduces a new measure of presence based on data that is unobtrusively obtained during the course of a VE experience. The measure is then applied in an experiment designed to explore the relationship between body movement and presence. It is shown that there is a positive association between these two, which is an important result for the design of interaction paradigms for immersive VEs.

2. Background

Presence research has focused on definition and ideas for measurement (Held and Durlach, 1992; Sheridan, 1996; Ellis, 1996; Slater and Wilbur, 1997) and there have been several empirical studies of contributing factors (Barfield et. al., 1995; Hendrix and Barfield, 1996a, 1996b; Welch et. al., 1996). Some of the factors studied have included the effect of visual display update rate, characteristics of the visual display system, the influence of spatialised sound, head-tracking, and interaction. Rather than attempting to directly measure presence, some studies (Pausch et. al., 1997) have examined the possible benefits of immersion on some aspect of task performance or behaviour within a VE. In this context 'immersion' means the objectively measurable characteristics of a system: the degree to which it delivers a surrounding environment accommodating many sensory modalities, with a high degree of match between sensory data and the participant's proprioceptive feedback. It could be argued that a central scientific goal of VE research is that of understanding the relative balance of 'immersive' factors that are necessary to generate a high degree of presence. The study of presence therefore is not just of academic interest, but such findings are important for the design of systems exploiting trade-offs between scarce computational resources.

This paper addresses the problem of how to operationalise and quantify presence. Most of the studies mentioned above use questionnaires, eliciting participant's subjective responses - of course always after the conclusion of the experience itself. While a useful starting point, the use of such questionnaires alone is never entirely satisfying, because of the high degree of difficulty in constructing scales that are truly comparable amongst different people and different applications.

Another approach is to attempt to measure presence by observation of people's behaviour. Held and Durlach (1992) suggested a 'startle' or looming response. This was extended by Sheridan (1992) to 'socially conditioned'

responses (would a person involuntarily put out their hand in response to a hand-shake gesture?). In (Slater, Usoh and Chrysanthou, 1995) an attempt was made to measure presence behaviourally by introducing contradictory information about an object represented in both the real and virtual world - with some information (visual) coming from the VE and other information (auditory) from the real world. The extent to which participants respond to the virtual information (allowing for differences in sensory preference) indicates their degree of presence in the virtual. A similar approach, though in the vestibular domain has been tried by Prothero (1995).

These behavioural techniques all suffer from the same problem: some feature has to be added to the environment (to cause the looming response, for example) that has nothing to do with the application, but is only there for the purpose of measuring presence. A good property of a measuring instrument should be the extent to which it can be used in any application without the addition of particular features that are for the sole purpose of measurement.

Figure 1 about here

In gestalt psychology (Kohler, 1959) there is the notion of figure and ground; within a single figure (e.g., Figure 1a) one aspect might come to the foreground, thus giving one interpretation, or another aspect might come to the foreground, resulting in a quite different interpretation. Presence in a VE occurs when the sensory data from the VE becomes the figure, with the surrounding real world forming the (back)ground. Just as in gestalt psychology it has been noted that transitions occur between figure and ground, so in VE experiences people often report such transitions between the 'real' and the 'virtual'.

Imagine then that it were possible to connect someone experiencing a VE to a gadget that gave a continual readout of their state of 'presence'. This gadget would have the same impact on understanding of VEs as the discovery of rapid eye movement (REM) sleep had on understanding of dreams and sleep (Aserinsky and Kleitman, 1953). Unfortunately, such a gadget has not yet been invented. In this paper, however, a simple method for monitoring presence during the course of a VE experience is introduced, exploiting the notion of transitions. It is a measure that is very simple to administer, and is applied during the course of the VE experience itself without actually disturbing the presence of the participant.

The new technique is introduced in the next section. It is applied in an experiment in Section 4. The purpose of the experiment was two-fold: to assess whether the new measure gives results that are comparable to the usual questionnaire results, and to examine a hypothesis that the degree of body movement required in a VE task is positively associated with presence. The results are given in Section 5, with discussion and critique in Section 6. The conclusions and some recommendations for future uses of the method are presented in Section 7.

3. A Presence Counter

Figures 1a-d are examples of the well-known result that the same information can be perceptually interpreted as quite different entities by the same person at different moments in time. Figure 1c for example will usually be first interpreted as three narrow triangular sectors radiating from the centre of the circle. However, after staring at

the centre for a while, the figure will suddenly reorganise itself into something different, and then every so often a spontaneous change from one interpretation to the other will occur (Kohler, 1959).

While in an immersive VE the participant receives a continuous stream of sensory data - mainly visual from the VE, but also often auditory from the real world, and of course real-world tactile and kinesthetic data (e.g., the weight of the helmet). Assuming that there are two alternate gestalts available (V: 'I am in the world depicted by the VE system', R: 'I am in a lab in the Computer Science building, wearing a helmet...') - at any moment of time the individual will favour one rather than another. Presence can be thought of as the extent to which the interpretation V is favoured. From this standpoint post-experimental questionnaire responses on presence can be considered as the individual's overall estimate of the proportion of time that they spent in state V.

During the course of a VE experience, an individual will typically experience transitions between V and R. The moments in time at which the individual switches from one interpretation to the other, in particular from V to R, are of particular interest. If it could be known when and why these occurred, this would be a major contribution to the problem of eliciting the factors that enhance or inhibit presence. The participants cannot be asked to report transitions from R to V, since this would require them to immediately break out of their state of presence in order to report back to the 'real world'. However, they can be asked to report transitions from V to R - at such a moment they have just 'returned' to the real world, realise that they are in an experimental situation in a virtual reality laboratory, and with that comes the memory that they are supposed to report such realisations.

This is analogous to the situation in the study of dreams. A researcher in a Dream Research Laboratory knows the likely onset of a dream by observation of the REM monitor. At any moment during the REM phase the sleeper can be awoken and asked to report the dream, or whatever other data is required[1]. In the case of the VE experience, if the state of presence is considered as equivalent to a dream, the dreamer is 'awakened' by whatever caused the break in presence. At that moment a report can be given that a 'break' has occurred without this in itself disturbing the sense of presence - which of course has already been disturbed.

Consider the following scenario: an individual enters a VE with the instruction to report whenever a 'break in presence' (BIP) occurs, and only at such a moment, to report this. At the end of the experience, lasting time t, there will be b such BIPs at times $t_1, t_2, ..., t_b$. The problem is to now use this information to recover the tendency of the individual to be in the 'presence' state (p), and also to understand the reasons why the BIPs occurred when they did. Here p is given a specific interpretation as the asymptotic (long term equilibrium) probability of being in state V.

There is clearly a difficulty in recovering p from the time sequence, since only half the information is available - i.e., when and how many times there was a break in presence is known, but the times when the individual entered the presence state are unknown. When b=0 (no transitions), for example, is this because the individual spent the

---

1. An excellent popular account of this type of research can be found in S. LaBerge, Lucid Dreaming, Ballantine Books, NY, 1985.

whole time present (in state V), or the whole time in state R? For any given value of b, there are two extreme interpretations - one where the unknown time is assumed to be in the V state, and the other when the unknown time is assumed to be in the R state. The discrepancy between these two interpretations decreases with increasing b, and it is easy to show that the expected value of p for increasing b is 0.5. (Appendix A).

p can be reconstructed with some simplifying assumptions. Suppose, as a simplifying assumption, that presence is binary - i.e., at any moment of time the participant is either in state R or state V. Divide the total time into n equal intervals (t=1,2,...,n). Denote by p_{ij} the probability that if at time interval t the participant is in state i then they will be in state j at the next time interval t+1 (i.e., a BIP has occurred on the boundary between these two time intervals). Here state 0 corresponds to R and state 1 to V. Note that this assumes p_{ij} to be independent of t, so the transition matrix

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$$

represents a stochastic process modelled by a two-state Markov chain (Karlin, 1969).

It is not difficult to show that $\mathbf{P}^k$ is the k-step transition matrix, its elements $p_{ij}^{(k)}$ are the probabilities that if at time t the individual is in state i then at time t+k they will be in state j. As k→∞ the equilibrium probabilities p_0 and p_1 are obtained, denoting the probabilities of being in the corresponding states in the long run (which given the assumptions of a Markov chain are independent of the initial state). A fundamental limit theorem of Markov chains shows that (in the particular case of the two-state chain):

$$p_0 = p_0 p_{00} + p_1 p_{10}$$
$$p_1 = p_0 p_{01} + p_1 p_{11}$$
$$p_0 + p_1 = 1$$

and therefore

$$p_0 = \frac{p_{10}}{p_{01} + p_{10}}$$

$$p_1 = \frac{p_{01}}{p_{01} + p_{10}}$$

The unknown p is interpreted as p_1, the equilibrium probability of being in state V.

The goal now is to use the observed data t_1, t_2, ..., t_b to estimate the transition probabilities p_{ij} under each of two alternate conditions, the first assuming a low propensity to presence and the second a high propensity (bearing in mind the two possible interpretations of b=0). In each case the t_i are assigned to the appropriate intervals, and mark a transition from state 1(V)to 0(R), such transitions assumed to occur at the boundary between the two intervals.

Low Presence Condition

There are b observed transitions from V to R (BIPs). If t is an interval at which there was such a transition, then at interval t-1 the participant must have been in state V. There are therefore b intervals with state V. It is assumed, for the moment, that intervals are small enough so that no two successive states reported a BIP, and that the first and last intervals did not report a BIP. *In the low presence condition, it is assumed that all intervals in which the state is unknown are in state 0.*

p_{00} is the proportion of times that an interval in state 0 is followed by an interval also in state 0. There are n-1-b intervals in state 0 that are followed by another interval (interval n has no successor). All but b of them are followed by intervals in state 0. Therefore

$$p_{00} = \frac{n - 1 - 2b}{n - 1 - b} \text{ and } p_{01} = \frac{b}{n - 1 - b}, \ 2b \leq \text{n-1.}$$

p_{10} is the proportion of times that an interval in state 1 is followed by an interval in state 0. Since this always occurs,

$$p_{10} = 1 \text{ and } p_{11} = 0.$$

From these the equilibrium probabilities are:

$$p_0 = \frac{n - 1 - b}{n - 1} \text{ and } p_1 = \frac{b}{n - 1}, \ 2b \leq \text{n-1.}$$

High Presence Condition

In this case the assumption is that all intervals in which the state is unknown are in the state V(1). A similar analysis to that above yields:

$$p_0 = \frac{b}{n - 1} \text{ and } p_1 = \frac{n - 1 - b}{n - 1}, \ 2b \leq \text{n-1.}$$

Let $p_C(b)$ be the equilibrium probability of being in state V with b BIPs observed and with C corresponding to the Low Presence (L) condition, or the High Presence (H) condition, then:

$$p_L(b) = \frac{b}{n-1}$$

$$p_H(b) = \frac{n-1-b}{n-1}$$

with $2b \leq n-1$.

Figure 2 about here

The relationship is illustrated in Figure 2, showing that when the number of BIPs achieves its maximum $(n-1)/2$, $p=0.5$.

Figure 2 highlights a problem: in practice only b is observed, and for any level of b there are two extreme values of p. Knowing only b gives insufficient information to estimate p. To choose between $p_L(b)$ and $p_H(b)$, therefore, a *discriminator* is required- some additional information to select one of the two alternatives. The simplest way to achieve this is a question at the end of the session, asking the participant to classify their overall experience with respect to their sense of presence. The answer to this together with the value of b would then allow an estimate of p.

It should be recalled that this analysis gives, for any condition, two extreme interpretations; for example, in the low presence condition, the analysis really implies:

$$\frac{b}{n-1} \leq p < \frac{1}{2}$$

In the absence of prior information it would be normal to use an estimate half-way between these two bounds. Such estimates would be linear transforms of $p_L(b)$ and $p_H(b)$, and therefore would have no effect on relationships with other variables discovered in statistical analysis. Therefore this paper continues to use $p_L(b)$ and $p_H(b)$, which should properly be referenced as 'extremal probabilities', although the qualifier 'extremal' is usually dropped.

For any choice of time interval, there can always be a situation where successive intervals report a BIP, or where there is a BIP in the first interval or in the last interval. It is important to be able to cater for these special cases, in order to avoid the problem of having to choose very large values of n, thus forcing the probabilities to the extremes.

The analysis can be easily adjusted to take account of this. One additional assumption is made, which is that if there are successive BIPs then the amount of time in the V state in between them is negligible and can be ignored.

Suppose that k out of the b BIPs are followed by a BIP in the next interval. Then the transition matrix probabilities are as follows:

$$\mathbf{P}\text{ (low presence)} = \begin{pmatrix} \dfrac{n - 1 - 2(b - k)}{n - 1 - b + k} & \dfrac{b - k}{n - 1 - b + k} \\ 1 & 0 \end{pmatrix}$$

with $p_L(b) = \dfrac{b - k}{n - 1}$.

$$\mathbf{P}\text{(high presence)} = \begin{pmatrix} \dfrac{k}{b} & \dfrac{b - k}{b} \\ \dfrac{b - k}{n - 1 - b} & \dfrac{n - 1 - 2b + k}{n - 1 - b} \end{pmatrix}$$

with $p_H(b) = \dfrac{n - 1 - b}{n - 1}$.

A further refinement allows for a BIP in the first or last intervals. Let $s_1 = 1$ if there is a BIP in the first interval, and 0 otherwise, and similarly $s_n = 1$ if there is a BIP in the last interval, and 0 otherwise. Then, following the same reasoning above it can be shown that the transition matrix probabilities are:

$$\mathbf{P}\text{ (low presence)} = \begin{pmatrix} \dfrac{n - 1 - 2(b - s_1 - k)}{n - 1 - (b - s_1) + k} & \dfrac{b - s_1 - k}{n - 1 - (b - s_1) + k} \\ 1 & 0 \end{pmatrix}$$

$$\mathbf{P}\text{(high presence)} = \begin{pmatrix} \dfrac{k}{b - s_n} & \dfrac{b - s_n - k}{b - s_n} \\ \dfrac{b - s_1 - k}{n - 1 - (b - s_n)} & \dfrac{n - 1 + k - 2b + s_1 + s_n}{n - 1 - (b - s_n)} \end{pmatrix}$$

Clearly b>k, and when b=0 then k=0 and $s_1 = s_n = 0$ and the probabilities would be 0 or 1 in this case.

Figure 3 about here

Figure 3 illustrates the ideas of this section. It shows the occurrences of the BIPs, and the corresponding low and high extremal probabilities (multiplied by 10) for a particular individual selected from the subjects in the experiment described in Section 4. The BIPs are taken as occurring on the boundaries between intervals, but actually occur within an interval. This is illustrated in the graph by the steeply sloping lines in an interval where a BIP has occurred.

4. Experiment

In this section an experiment is described that had two principal goals. First, the above technique is used to estimate p. If the approach is sound then there should be a significant positive correlation between p and post-experimental questionnaire results relating to presence - following the hypothesis that these questions are answered on the basis of the balance of time that a participant spends in the V compared to the R state. Second, previous results (Slater, Usoh, Steed, 1995; Slater, et. al., 1997) suggest a positive relationship between the degree of body movement of a participant and their reported and behavioural presence. The experiment also examines this idea in a different context to those previous studies.

There was also a subsidiary goal. Elsewhere the positive benefits of beginning a VE session in a virtual environment that itself depicts the real laboratory in which the experience is taking place have been noted (Slater, et. al., 1997). This allows a degree of adaptation and training in this so-called 'virtual ante-room' before the main VE experience is started. This experiment examines whether the actual initial virtual place itself matters, or whether it is just something *different* from the main VE scenario.

The experimental scenario involved the participants observing a sequence of moves on a Tri-dimensional chess board (as introduced by the Star Trek TV series). This was chosen because it is a quite large and complex three-dimensional object, and fitted well with the requirement to induce significant body movement in participants who were required to reach out and touch the chess pieces. A stereo pair of the Tri-dimensional chess board is shown in Plate 1. As can be seen it is a structure with several layers, resting on a table. The dimensions are:

Object	Dimension (m)
Table top	0.74
Large chess boards	0.2×0.2
Small chess boards	0.1×0.1
Lowest board: height above table top	0.22
Middle board: height above table top	0.42
Highest board: height above table top	0.62
Small boards: height above large board	0.1

The pieces that had to be touched were distributed over the entire board. Hence the highest piece was about 1.46m above the ground - which for some subjects required considerable stretching to be reached. It is vital to note that the measure of body movement used, namely the total amount of hand movement, is of course, a measure of whole body movement - since it would encompass such reaching and stretching.

The factorial design was for 20 participants divided into 4 groups, as shown in Table 1.

Table 1

Factorial Design

Place:	Low Activity	High Activity
Different From Lab	5	5
Same as Lab	5	5

Each subject was paid £5 for completing the study. In the event two subjects were unable to either understand or properly follow the instructions and were replaced by two other subjects, so 22 people completed the experiment, with 2 cases discarded. In the final 20 there were 18 men and 2 women, who were recruited through advertisement on the campus. There were 5 undergraduates, 6 Masters Students, 4 PhD students, 2 Research Assistants, 1 Faculty, and 2 miscellaneous others. No subject had anything at all to do with the research itself.

Each participant started the experience in a virtual room (the virtual ante-room). This was either a depiction of the real laboratory in which the experiment was taking place ('Same as Lab') or another room of the same size but different furniture and layout ('Different from Lab'). After receiving instructions the participants made their way through a door to a field with trees and plants outside. Some 5 metres beyond the door was a table with a Tri-Dimensional Chess board (Plates 2 and 3). Those assigned to the 'Low Activity' group were told to repeatedly look for a red chess piece, and when found press a button on a 3D mouse that they were holding throughout, and observe the movement of the piece. Those assigned to the 'High Activity' group were told that when they observed the red piece to reach out and touch it, and it would then move. At the end of an entire sequence of 9 moves, a large button on the side of the table would turn red. The 'Low Activity' group had to click the button on their hand-held 3D mouse and the 'High Activity' group had to touch the virtual red button. All participants were told that when they noticed that the sky had become dark, they should return from the field to the starting room. The sky was darkened after three complete sequences of moves, and the mean and standard deviation of the time spent in the field was 319±64s. All subjects were told that they would be asked about the sequence of moves observed after the experience.

Prior to starting the experiment all subjects were asked to complete a short questionnaire that obtained background information - gender, job status, and prior experience of virtual reality. Also some information about their pattern of memory of any other place that they had visited earlier in the day was required (information which has not been analysed at the time of writing this paper).

After completing the questionnaire, the subject was shown each of the Figures 1b-d in turn. (Most subjects had seen Figure 1b before, so that this was hardly used). They were asked to describe their initial interpretation of the figure. (For example, for Figure 1c, most saw the three thin triangles first of all). Then they were asked to stare at the figure, and notice if any change occurred. If a change did happen, they were asked to continue to observe the figure and clearly exclaim 'Now!' if and whenever it spontaneously reconfigured itself to look the same way as when they first saw it. After this training they were given the following instructions to read:

IMPORTANT: Transition to Reality

When you enter the virtual reality you may have the sense of being in 'another place'.

In just the same way as you saw transitions in the images that you just looked at, you may experience transitions in your sense of place -

Virtual: sometimes you will be in the virtual place

Real: sometimes you will become aware of the real lab in which the experience is really happening.

If and only whenever you experience a ***transition to Real***, please say *'Now'* very clearly and distinctively.

All instructions were reinforced verbally once the subject entered the real laboratory itself, and then again while they were in the 'virtual ante-room' before entering the field with the chess board. While in the ante-room they were shown how to move around, how to make a small red cube on a table respond by either touching it ('High Activity' group) or by clicking with their forefinger on the 3D mouse ('Low Activity' group).

The virtual reality laboratory is in a small enclosed room within a large laboratory, in which there is continual noise (constant noise of workstations, and random noise of occasional phone rings or conversations). No attempt was made to reduce background noise or further isolate the VR room from the remainder of the laboratory: indeed there was interest as to whether the background events would trigger transitions from V to R.

The scenarios were implemented on a Silicon Graphics Onyx with twin 196 MHz R10000, Infinite Reality Graphics and 64M main memory. The software used was Division's dVS and dVISE 3.1.2. The tracking system has two Polhemus Fastraks, one for the HMD and another for a 5 button 3D mouse. The helmet was a Virtual Research VR4 which has a resolution of 742×230 pixels for each eye, 170,660 colour elements and a field-of-view 67 degrees diagonal at 85% overlap.

The total scene consisted of 13298 polygons which ran at a frame rate of no less than 20 Hz in stereo. The latency was approximately 120 ms.

Subjects moved through the environment in gaze direction at constant velocity by pressing a thumb button on the 3D mouse. They had a simple inverse kinematic virtual body (Plate 4) . When they reached forward to touch a chess piece they would see their virtual arm and hand.

At the end of the session subjects were given a second questionnaire. The main purpose of this was to gather information on their sense of presence. There was an initial question that asked for the reason why (if this was the case) they reported no or very few transitions, giving four options: rarely being in the virtual world, almost always being in the virtual world, forgetting to report transitions, other reasons. In retrospect this question was not particularly useful, since it only required an answer when subjects reported 'no or very few transitions' without giving a definition of this. No subject reported 'forgetting' to report transitions.

A second question was open ended, asking for the 'causes of the transitions' (whether or not these had been reported at the time). There were five questions relating to presence interspersed through the questionnaire, each rated on a 1 to 7 scale, where 1 indicated low and 7 high presence. The first was a priori considered the most direct elicitation of presence and used as the discriminator: a score of more than 4 on this resulted in the formula $p_H(b)$ being used, otherwise $p_L(b)$.

1. Please rate *your sense of being in the field*, on the following scale from 1 to 7, where 7 represents your *normal experience of being in a place*.

I had a sense of "being there" in the field:

1. Not at all ... 7. Very much.

2. To what extent were there times during the experience when the field became the "reality" for you, and you almost forgot about the "real world" of the laboratory in which the whole experience was really taking place?

There were times during the experience when the virtual field became more real for me compared to the "real world"...

1. At no time ... 7. Almost all the time.

3. When you think back about your experience, do you think of the field more as images that you saw, or more as somewhere that you visited ? Please answer on the following 1 to 7 scale:

The virtual field seems to me to be more like...

1. images that I saw ...7. somewhere that I visited.

4. During the time of the experience, which was strongest on the whole, your sense of being in the field, or of being in the real world of the laboratory?

I had a stronger sense of being in...

1. the real world of the laboratory ... 7. the virtual reality of the field of plants.

5. During the time of the experience, did you often think to yourself that you were actually just standing in an office wearing a helmet or did the field overwhelm you?

During the experience I often thought that I was really standing in the lab wearing a helmet....

1. most of the time I realised I was in the lab ... 7. never because the virtual field overwhelmed me.

Some data were automatically collected during the course of the experiment - in particular the times at which the participant said 'Now!', the total time in the virtual field, and the amount of hand and head movement.

5. Results

General

The overall levels of reported 'presence' as ascertained from the questionnaire responses were high. Figure 4 shows the median response for each for each of the five presence related questions, showing, for example, that on question 1 (the discriminator) half of the responses were at level 5 or more.

The number of BIPs ranged between 0 and 14. The mean time between BIPs was 48±37s, the minimum time interval was 5.5s and the maximum 141s. The time interval used for the analysis was 10s, this being approximately the largest compatible with the assumptions that $2b \leq n-1$. There were two cases where there were some BIPs in sequence, and two other cases where there was a BIP in the first or last interval.

Figure 4 about here

Relationship between p and Questionnaire Based Presence

The first question to consider is whether there is a relationship between the estimate of presence p, and the presence questionnaire responses. The usual approach of the authors to combining the results of the presence questions into one overall score (without resorting to averaging across ordinal data) is to count the number of high scores ('6' or '7'), thus giving each subject a count out of 4, *for the questions other than that used as the discriminator* (Appendix B).

Figure 5 about here

Figure 5 shows the scatter plot of p against the questionnaire presence count (recall that the p's are 'extremal'). There is a clear positive relationship, though with an outlying point. Even including this point there is a statistically significant correlation between p and the presence count ($r^2 = 0.32$, $t = 2.920$, $t_{18} = 2.101$ at 5% significance level). When this outlier is removed the result improves substantially ($r^2 = 0.65$, $t = 5.588$, $t_{17} = 3.965$ at 0.1%). Examining the responses of the particular person involved, he wrote that he was disturbed by the absence of sound in the VE, knew that the experimenters were in the real lab alongside him, and that he wanted to talk to them because exploring an environment is often a "communal activity". The experimenters' notes record that he did indeed continue to talk to them during the immersive experience. He gave a '3' response to the discriminator question (writing "SOUND!" next to his response), and scores of '7' for each of the remaining four presence-related questions. The assignment of this person to the low presence condition on the basis of this particular discriminator is therefore dubious.

Relationship between p and Hand Activity

The next issue to consider is the relationship between p and the main independent factors - Place (the ante-room model) and Activity. The type of place depicted in the virtual ante-room had no significant influence on presence, and this factor is ignored in further discussion. Table 3 shows the means and standard deviations of p for the activity groups, and the difference in means is not significant although contrary to expectation, the low activity group seems to have a higher 'average presence' than the high activity group. There is a highly significant difference in variance (the variance ratio is 14.0, $F_{9,9} = 3.2$ at 5%), with much less variation amongst the low activity group. The difference in variation was to be expected, since the inactive group were not required to move their hand at all (except pressing the button on the hand-held 3D mouse for navigating. This in fact did not require any movement of the hand relative to the body to accomplish).

Table 3

Means and Standard Deviations for p by Activity

	Mean and Standard Deviation
Activity High	0.60 ± 0.44 (n=10)
Activity Low	0.89 ± 0.12 (n=10)
	$t = 1.99$ ($t_{16} = 2.120$ at 5%)

A more detailed examination reveals a different picture. Figure 6 shows a scatter plot of p by total hand movement per unit time, discriminating between the two activity groups. It suggests that there is no discernible relationship between the amount of hand movement of the low activity group and p, as would be expected. However, for the high activity group there appears to be a positive linear relationship between hand movement and p.

Figure 6 about here

Table 4(a) gives the result of a multiple regression analysis of p on Activity and hand movement (hm), allowing for the possibility of differing slopes for the low and high activity groups. There is a significant positive slope for the high activity group, with an overall squared multiple correlation of 0.42.

Table 4

Regression of p on activity and hand movement per unit time (hm).
Allows for different slope for hm for each level of activity.

(a) Based on n = 20 observations,
with d.f. = 16, t_{16} = 2.120 at 5%.

Overall R^2 = 0.38

Activity	Regression	t for slope
High	p = -0.91 + 11.07 hm (4.90)	2.26
Low	p = 0.87 + 0.34 hm (4.78)	0.07 (N.S.)

(b) Based on n = 19 observations,
with d.f. = 15, t_{16} = 2.131 at 5%.

Overall R^2 = 0.73

Activity	Regression	t for slope
High	p = -1.90 + 19.11 hm (3.36)	5.69
Low	p = 0.87 + 0.34 hm (2.92)	0.11 (N.S.)

Inspection of Figure 5 shows an outlying point - with a low presence and high hand movement (above 0.15 mps) in the high activity group. This was caused by the same person who was the outlier in Figure 4.

Removing the data for this person from the analysis results in Table 4(b). The multiple correlation increases to 0.78, and the slope for the high activity group is well into the highly significant range.

Explanations for BIPs

A question asked the participants to give the reasons for their transitions to the real:

If you did make transitions from virtual to real, whether or not you reported these at the time, what do you remember as the causes of the transitions? (For example, hearing an unexpected noise from the real lab might cause such a transition).

The reasons given can be classified into two main (most often reported) types:

External: Sensory information from the real world intruded into or contradicted the virtual world, either in the form of noises or people talking, or else the touch or feel of interactions with real solid objects (such as the virtual reality equipment itself).

Internal: This is where something 'wrong' with the virtual world itself is noticed: for example, the laws of physics not being obeyed, objects looking unreal, the absence of sounds, display lag.

There were a number of subsidiary reasons:

Experiment: Some aspect of the experimental set-up itself, or the instructions intruded.

Personal: Some personal feeling intruding, such as embarrassment or consciousness of being observed from the outside.

Attention: A loss of attention to what is happening in the virtual world, or some aspect of the virtual world that results in a loss of presence.

Spontaneous: A BIP for no (conscious) apparent reason.

Table 5 gives the number of participants who responded in each of these categories, and some examples of each.

Table 5

Reasons Given for Transitions to the Real

n is number of participants who gave responses in the corresponding category

Cause	n	Some Examples
External sound	7	"Noises from the lab (people talking)."
		"Hearing background noise."
External touch or force	9	"I was supposed to be in a grass plane, but when I moved my feet I realised it was a plank under my feet (in the real)."
		"Feeling of the floor under my feet."
		"Becoming aware of cable wrapped around foot."
		"The cable brushing against my legs."
		"Trapped in wires."
Internal	11	"The length of time taken to interact with the world."
		"Turning and thus becoming aware that the virtual world was not the real world."
		"If moved head quickly."
		"The time taken for the chess pieces to move."
		"The way the sky darkened, not smooth, like someone had switched off the sun."
		"Weird things happening that are obviously not real (eg, the chess set)."
		"The more I needed to examine the contents of the 'virtual' the more my awareness flipped into the 'real'."
		"Became very close to the chess board."
		"Getting too near to things (especially trees)."

Experiment	3	"Having to remember the real world instructions."
		"The task of being asked to monitor the changes from virtual to reality itself creates a sense of going back to reality."
		"Once experienced a transition, became sensitive to it happening again."
Personal	3	"Wanted to talk to the experimenters to share the experience."
		"Embarrassment."
		"Very conscious."
Attention	3	"Transitions occurred between the tasks, for example, when looking for the next red piece, but only if I couldn't see it at first glance."
		"Attention wandered after realising that the chess sequence was iterative."
		"Not having a task to do."
Spontaneous	2	Spontaneous feeling.
		It just occurred to me.

The Discriminator Question

The analysis above hinges on the choice of the 'discriminator' question, since this classifies each participant into a 'low' or 'high' presence group, and therefore determines the computation of p. A different discriminator question could lead to quite different results. However, the results for this experiment are robust with respect to choice of discriminator question.

The analysis was repeated for each of the remaining four presence-related questions, and also for the average of all of the five questions (Table 6). (Of course, the 'outlying point' corresponding to someone who had written '3' for question 1, but '7' for each of the others, does not occur for any of the other choices of discriminator). For every choice of discriminator question, except for question 3, the results are the same. When the mean response of each of the presence questions is used as discriminator, again the results are the same.

Table 6
Regression Analyses
p on Hand Movement (hm)
Computing p with Different Discriminators

Only the High Activity equations are shown (no Low Activity result is significant).
$n = 20$, d.f. $= 16$, $t_{16} = 2.120$

Question used as Discriminator	R^2	Regression: p =	t-value
1. Please rate *your sense of being in the field*, on the following scale from 1 to 7, where 7 represents your *normal experience of being in a place*.	0.38	-0.91 + 11.07 hm	2.26
2. To what extent were there times during the experience when the field became the "reality" for you, and you almost forgot about the "real world" of the laboratory in which the whole experience was really taking place?	0.42	-1.64 + 17.00 hm	3.41
3. When you think back about your experience, do you think of the field more as images that you saw, or more as somewhere that you visited ?	0.26	-0.23 + 7.78 hm	1.46 NS
4. During the time of the experience, which was strongest on the whole, your sense of being in the field, or of being in the real world of the laboratory?	0.77	-1.46 + 16.23 hm	7.01
5. During the time of the experience, did you often think to yourself that you were actually just standing in an office wearing a helmet or did the field overwhelm you?	0.43	-1.64 + 16.67 hm	3.20
Discriminator is computed as average of responses to questions 1-5.	0.77	-1.46 + 16.23 hm	7.01

(Note that question 4 has exactly the same impact as a discriminator as the average).

6. Discussion

The method presented in this paper relies on a number of assumptions.

1. Presence in the 'real' and 'virtual' is treated as a binary state.

The authors would not seek to defend this as a statement about the psychological processes involved. It is used here in the spirit of a simplifying assumption, to allow the construction of the stochastic model. It should be considered as a starting point, allowing the construction of first model of this type, and an attempt will be made to supersede this in later work.

2. The stochastic model assumes discrete time.

Again this is a simplifying assumption often employed in the initial stages of constructing a model of complex phenomena. It may be possible to employ a continuous time stochastic model instead.

3. The transitions can be modelled as a Markov chain.

This assumes that the transition probabilities are one-step - that what happens in any interval is statistically independent of all other intervals except for the last. The veracity of this assumption is unknown, and again should be viewed as a simplification for the purpose of the construction of an initial model.

4. The requirement to report BIPs does not in itself influence the participants to report BIPs.

There is experimental evidence to support the argument that the requirement to report BIPs increases the chance of BIPs occurring. Girgus, Rock and Egatz (1977) found that giving subjects a knowledge of the reversibility of ambiguous figures substantially increased the chance of these being reported. About half the subjects who were not told about the reversibility of figures never reported a transition, whereas all of the subjects who were told about the reversibility always reported transitions. This raises the possibility that more BIPs were reported than would otherwise have naturally occurred.

This is a difficult issue, since it could also be argued that the requirement to report BIPs sets up a dual task for the participants - to do their actual task in the VE and to pay attention to their state in order to be able to report the BIPs. A counter argument to this is that it can also be argued that the requirement to report a BIP might only enter consciousness at the time immediately after a BIP has occurred (because for the rest of the time the participant is in the V state).

A preferable response to these problems is to agree that the method for reporting BIPs, relying on a verbal response is certainly not an ideal way to obtain this information. It is an interesting and challenging research topic to try to find physiological correlates to BIPs that can be measured unobtrusively.

5. The discriminator question can discriminate between the low and high propensity cases.

Use of a discriminator question does indeed result in an uncomfortable reliance on questionnaire data. An alternative, behaviourally based discriminator, would be preferable.

7. Conclusions

Notwithstanding the critique above, a new method for measuring presence in virtual environments has been introduced, where the major component of the measure depends on data collected during the course of the VE experience itself. It is based on the number of transitions between a state of being in the VE, to the state of being in the real world. Using the simplifying assumption that the presence state is binary, and that changes in state between presence and non-presence form a time independent Markov chain, an equilibrium probability of presence can be computed. This requires only one additional post-experimental discriminator question concerning each participant's assessment of whether they had been in the presence state for more or for less than half the time.

This new technique was tried in an experiment to assess the impact of body movement on presence, and was found to be positively and significantly correlated with the usual questionnaire based measure. It is encouraging that the questionnaire score and the new measure were correlated. However, because the new measure is based on data collected during the course of the experience itself, and requires participants to respond only when they realise that they have exited the presence state, the authors place more reliance on the new approach than the questionnaire based approach. They would argue that for the first time there is a measure that is really eliciting something very close to the underlying phenomena itself.

The major issue of the experiment apart from the methodology for presence measurement, was the relationship between presence and hand movement. (In this experiment head and hand movement were significantly correlated, $r^2 = 0.54$, t = 4.6, $t_{18} = 2.101$ at 5%). The evidence strongly suggests a positive association between presence and hand movement, in line with previous evidence on the relationship between presence and body movement. The direction of causality is unknown, but the authors suspect that there is a two way relationship: high presence leads to greater body movement, and greater body movement reinforces high presence. A specific study to examine this is required.

Nevertheless, the data also suggests an 'average' presence for the 'low active' group that is higher (though not significantly so) than the active group. There is an explanation that might account for this. In the previous quoted studies where presence was associated with body movement, the active participants did not actually interact with objects. The tasks required them to move their bodies more (bend down, look up, walk in place) than the control group, but they did not touch any objects. In this experiment the active participants did interact and touch objects, whereas the 'low active' group did not. It is possible that the difference in mean presence score is because the 'high active' group had a greater opportunity to notice internal inconsistencies (no collision detection, close-up views of objects, for example) compared with the group who only watched without touching.

The relationship between presence and body movement follows from the notion that one of the most important determinants for presence is the requirement for a match between proprioception and sensory data. This is consequential for the design of interaction paradigms - where semantically appropriate body movement, exploiting proprioception, is preferred, for example, to the importation of techniques from 2D interfaces. This relationship between presence and proprioception in the design of interaction paradigms has been exploited by (Slater and Usoh, 1994; Mine, Brooks, Sequin, 1997).

There are some recommendations for future use of the new measure. First, the discriminator question should ask for the information that is required in a very direct manner. (In this sense question 4 would have been preferable to question 1). As soon as the VE experience has terminated, the participant should be asked to estimate the overall proportion of time spent in the presence state, crucially whether above or below 50 per cent. The exact wording of this discriminator question, or indeed as noted in Section 6, whether there is some better way to obtain this information should be given more thought.

Second, a standard time interval should be agreed, so that results can be easily compared between different applications and systems. In this experiment the time interval was chosen to be the greatest compatible with the requirement that $2b \leq n-1$ (n is the number of intervals). The interval used was 10 seconds. The choice of large values of n grants undue weight to the statistical significance of the count data, and pushes the probability estimates to more extreme values (though does not alter the relationship between them).

It should be noted, however, that the results are robust with respect to the range of time intervals. An analysis with intervals ranging from 1 second through to the maximum compatible with the crucial requirement of $2b \leq n-1$ always gives the same results. Although it seems reasonable to model a BIP as an instantaneous event, it is not obvious that this approach is suitable for the process of becoming present; the boundary from R to V being more fuzzily demarcated. Any agreed standard time interval should take this into account although a preferable solution would be to construct a model that did not require the use of discrete time intervals.

Finally, to return to the issue of body movement. In the Tri-dimensional chess experiment, it is clear that many of the (active) subjects *are learning about the chess board with their whole bodies.* As stated earlier, to call the movements 'hand movements' is an understatement of what is being measured. It is not surprising that amongst the active group those who exhibited greater body movement also tended to have a greater sense of 'being' in the same space as the chess boards. The question of the balance between the effect of body movement compared to the greater possibility of becoming aware of problems within the VE remains open. This is an opportunity to exploit Ellis' (1996) idea of iso-presence curves clearly showing trade-offs between different factors.

Appendix A

The Expected Time 'Present' as the number of BIPs increases.

Consider the transitions from R to V and V to R as occurring at instantaneously at random moments in time according to a Poisson process - that is, each time interval is an independent exponentially distributed random variable. The result is the same and mathematically simpler if everything is normalised by the total time, leading to 2b observations from the uniform probability distribution on the interval [0,1]. Suppose the BIPs occur at times $t_2, t_4, ..., t_{2b}$. The times at which the transitions R to V occur are $t_1, t_3, ..., t_{2b-1}$, with $0 < t_1 < t_2 < ... < t_{2b}$. Therefore, the total time in the state V, T_V is given by the formula:

$$T_V = \sum_{i=1}^{i=b} \left(t_{2i} - t_{2i-1} \right).$$

The expected value of t_i is $i/(2b+1)$, resulting in

$$E(T_V) = \frac{b}{2b+1}$$

As $b \rightarrow \infty$, $E(T_V) \rightarrow \frac{1}{2}$.

In the time interval from t_{2b} to 1, there may be another transition from R to V. However, the expectation of this extra time in V will tend to zero with increasing b.

(a)

(b)

(c)

(d)

Figure 1
Gestalt Images

Figure 2
Relationship Between Number of BIPs and Overall Presence p

Figure 3

A Time Sequence Showing Occurrences of BIPs
and the Corresponding Probabilities

Figure 4
Median Levels of Reported Presence

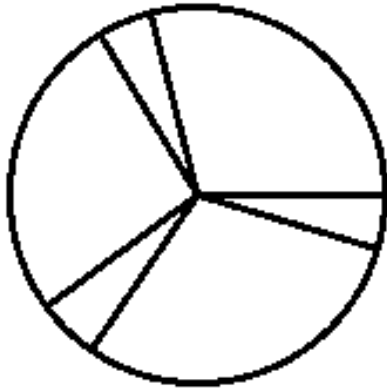**Figure 5**
Scatter Plot of p against Number of High Scores

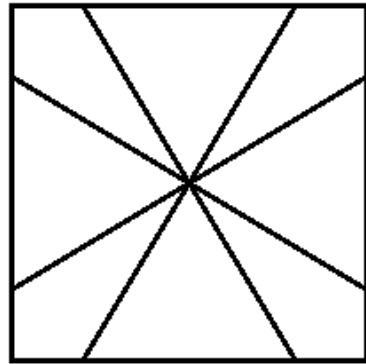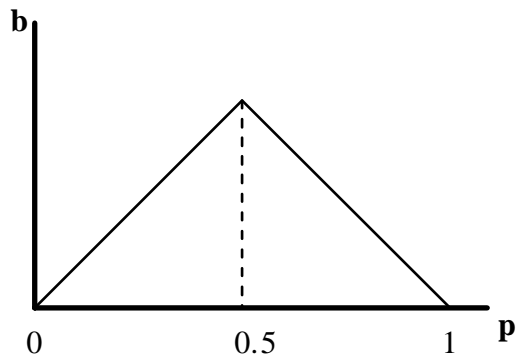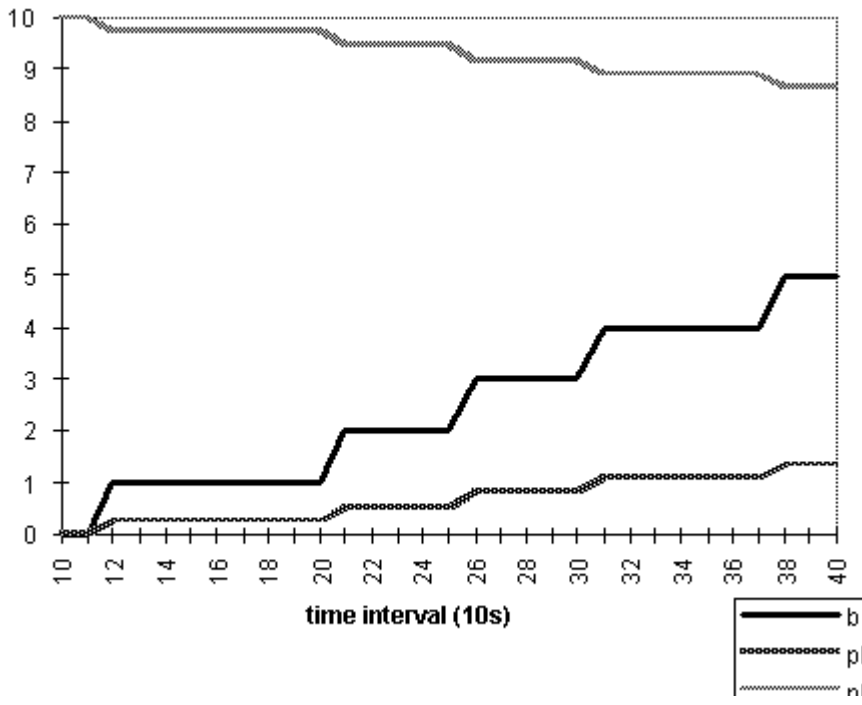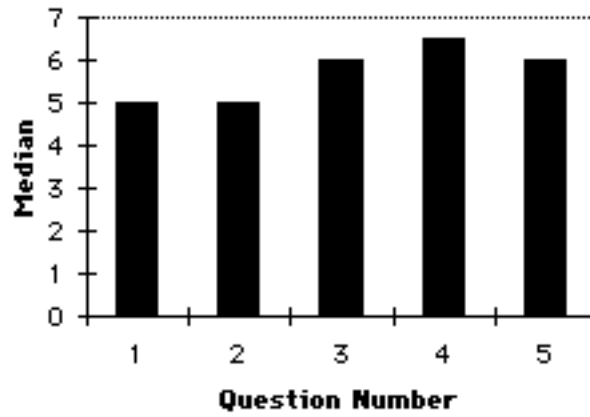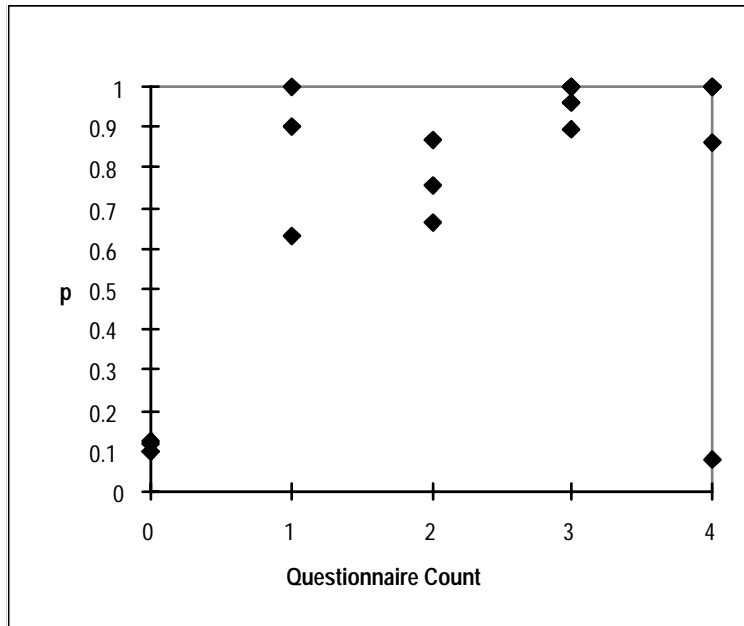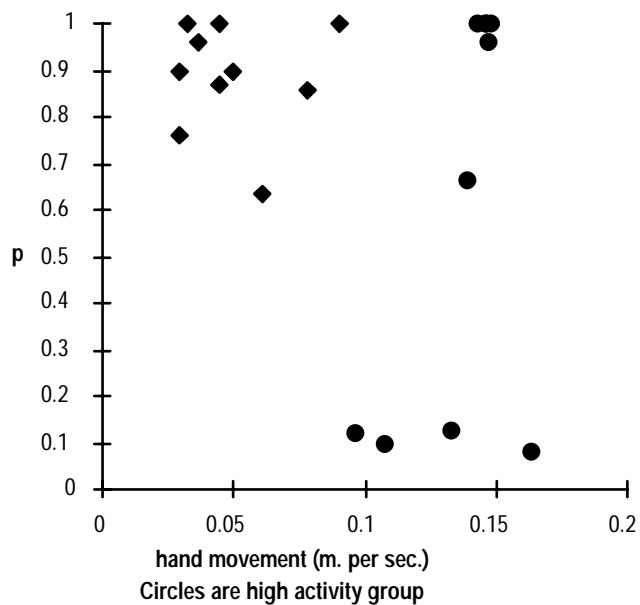**Figure 6**
Presence against hand movement

CHAPTER 10   *Being Together Through Virtual Touch*

1. Introduction

This paper describes an ongoing experiment to study whether haptic communication through force feedback can facilitate a sense of togetherness between two people at different locations while interacting with the same virtual environment (Durlach and Slater, 1998).

The experiment concerns a scenario where two or more people are at remote sites, but must co-operate to perform a joint task or play a game in a shared VE. In the current experiment, the set-up is an abstraction from a real situation, in order to simplify the interactions that occur in real life and to create a more controlled context suitable for an experimental study in the laboratory. We focus mainly on the impact of haptic display on the perceived quality of the interaction itself.

The sense of *presence* of a person in a VE has been of increasing interest to researchers, as discussed in the companion paper. In addition, there have been several studies on the development of social relations in shared VEs, and also on task performance (Bowers, 1996; Schroeder, 1997). However, there has been little attention paid to co-presence, that is the sense that participants have of being with other people, and to our knowledge, no attention paid at all to what the addition of touch and force-feedback between people would contribute to the shared experience. In this regard, the purpose of the experiment was to assess the impact of force feedback in addition to visual display

- On performance of the task
- On the sense of being together as reported by the participants
- On the extent to which participants could make guesses about the 'personality characteristics' of one another based on what they could see and feel of the behaviour of the other person during the course of the experiment.

2. Haptic Feedback for Shared Virtual Environments and Teleoperators

Haptic display of 3D objects in virtual environments has been a growing research area for scientists and engineers during the last few years. (Refer to Srinivasan, 1995 and Srinivasan and Basdogan, 1997 for a brief review of the current literature and the summary of research status). Analogous to graphical rendering, haptic rendering is concerned with real-time display of the touch and feel of virtual objects to a human operator through a force reflecting device. Our group at MIT has developed efficient haptic rendering methods for displaying the shape and surface details of 3D polygonal objects in VEs (Ho et al., 1997).

Although haptic display of 3D objects is being developed for various applications such as medical simulation and computer-aided design, its applications for shared virtual environments and teleoperators has not been explored previously. It is likely that the addition of haptic modality to shared virtual environments having visual

and/or auditory displays may increase the sense of "being together" and the quality of the interactions while performing collaborative tasks.

3. Experiment

3.1 Components of the Experimental Set-Up

The experimental set-up includes a dual 300 MHz processor IBM compatible personal computer, running Open Inventor to display the graphical model of the 3D virtual environment, and a force feedback device, PHANToM (SensAble Technologies Inc.), to convey to the user a sense of touch and feel of virtual objects.

3.2 Design

During the experiment, subjects were asked to play a collaborative game in virtual environments. They played the game with one of the experimenters who was an "expert" player. The subject was not allowed to know the "expert" player, and had no idea where the "expert" player was located during the experiment. The players were in different locations but saw a common scene and could feel the objects in the scene. The shared visual scene included a ring, a wire, and two cursors (green and blue small spheres that represented the contact points) attached to the ring (Figure 1). They were asked to move a ring on a wire in collaboration with each other such that contact between the wire and the ring was minimised or avoided.

Each subject manipulated his/her own cursor through a stylus attached to the force feedback device placed next to their seat. When the subject manipulated the stylus of the touch device with his/her right hand, the cursor moved in 3-D space, so that the ring could be moved.
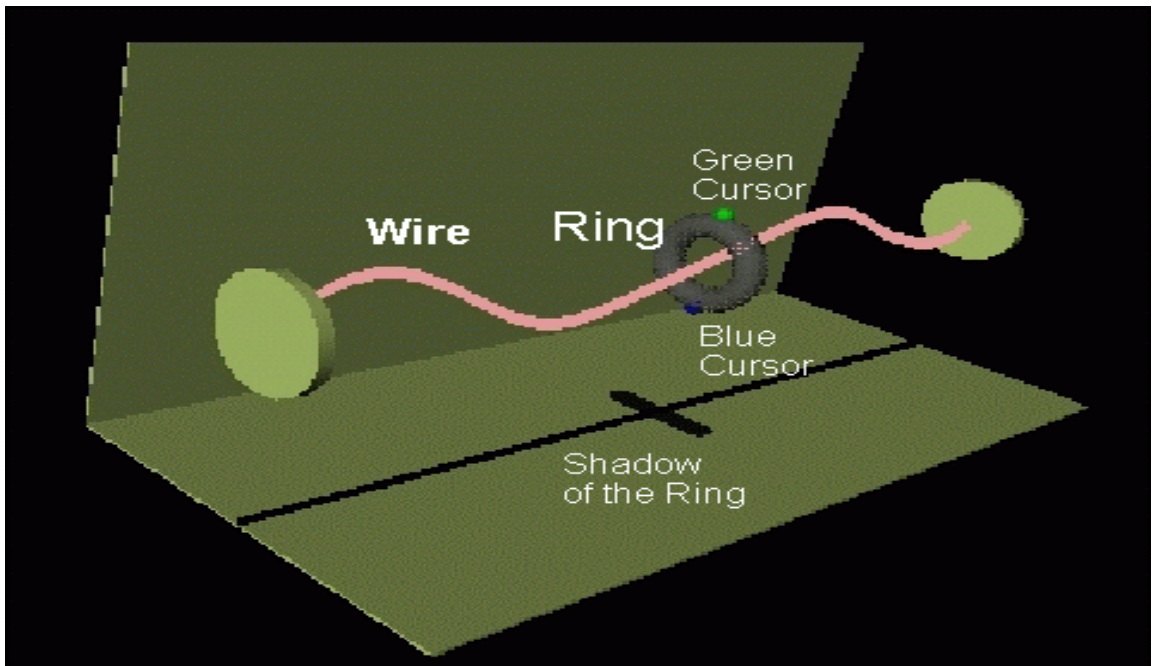
Figure 1. A shared virtual environment was created to play the "*Ring on a Wire*" game. Two subjects, represented by blue and green cursors and physically located in two separate rooms, share the same VE to move the ring in collaboration. Each subject can feel the resistive force through the haptic device when (1) the ring touches the wire (2) he/she pulls or pushes the other person.

The goal of the game was to move the ring with help from the other person without touching the wire. If the ring touched the wire, the colors of the ring and the surrounding walls changed to red to warn the subject of an error. They changed back to their original colors when the subjects corrected the position of the ring. To hold the ring, both subjects needed to press on the ring towards each other above a threshold force. If they did not press on the ring at the same time, the ring did not move and its color changed to gray to warn them. To move the ring along the wire, they each needed to apply an additional lateral force. Moreover, the shadow of the ring was displayed on the ground to give cues to the subject about its position relative to the wire.

The subjects were asked to move the ring back and forth along the wavy wire 3 times per trial. The shape of the wire was changed each time they reached the target end of the wire. The subjects could feel the forces through the haptic display (1) when the ring touched the wire (2) when a movement was induced by the other subject.

At the time of writing, we have tested 10 subjects. Five of these experienced first the haptic and visual condition, and then between 11 and 15 days later, the same scenario but without any haptic feedback at all. The second group of five subjects experienced the visual only condition. It is intended that they will later carry out the tasks again but with the haptic feedback enabled. Each pair participated in at least 10 trials which took about 30 minutes.

3.3 Variables Measured

We aimed to understand the effect of haptic cues on the sense of being together in VEs using *subjective* measures (see the descriptions of these terms in the companion paper by Durlach and Slater, 1998). After the experiment, each subject answered a questionnaire, which supplied basic demographic and background information. Subjective questions were asked in four categories including their (1) performance, (2) their sense of 'being together', (3) emotional reactions, and (4) personality profile.

Performance: Each person made a self-assessment of their own performance and the performance of the other person using the questionnaire. Sample questions in this category include:

Please give your assessment of how well you contributed to the successful performance of the task.

1. Not good at all ... 7. Excellent

Please give your assessment of how well you and the other person together performed the task.

1. Not good at all ... 7. Excellent

Sense of Being Together: Each of the following questions was rated on a 1-7 scale, where 7 meant a greater sense of 'togetherness'. There were seven questions in this category, sample questions include:

To what extent, if at all, did you have a sense of being with the other person?

1. Not at all ... 7. Very much so.

To what extent were there times, if at all, during which the computer interface seemed to vanish, and you were directly working with the other person?

1. At no time ... 7. Almost all the time

Emotional Reaction: To see if the experiment had any emotional impact on each subject, we included a few questions such as:

To what extent did you feel embarrassed, with respect to what you believed the other person might be thinking about you, in the way that you carried out this task?

1. Never ... 7. Almost all the time

Personality Profile: This area of study is new, and we wanted to 'push the limit' to examine whether it was possible to guess about the personality of the remote partner from these forms of interaction. We asked each individual to complete a standard personality profile test (Leary, 1983) supplemented by some additional questions particularly relevant to this task. We asked each person to complete this test for him/herself, and also to complete

the test guessing the answers for their remote partner. The purpose was to examine whether subjects' assessments of their unknown partner would change under the experimental conditions.

4. Results

At the time of writing we have an unbalanced design: a 'within subjects' experiment of five people (who experienced first visual and haptic, and then later visual only). A second group of five experienced the visual condition only. The analysis below is therefore in two parts - first comparing the within subjects experiment in itself, and second a 'between subjects' analysis comparing the visual plus haptic results of the first group of five, with the visual only results of the second group.

There were seven questions that related to the *sense of togetherness* experienced. A conservative measure of the overall level of togetherness experienced by a player is realised through counting the number of high scores (6 or 7) amongst the seven different questions. This is in line with scores used in previous studies of presence (Slater and Wilbur, 1997). A plot of this sense of togetherness against the maximum score realised in the game is shown in Figure 2 for the visual plus haptic group. The correlation $r = 0.98$ which is significantly different from zero at the 1% level. If mean game score is used instead of maximum score, then the result is similar, with $r = 0.88$ which is significant at 5%. While not too much can be made of this result, relating perceived 'togetherness' with task performance, this is at least encouraging.
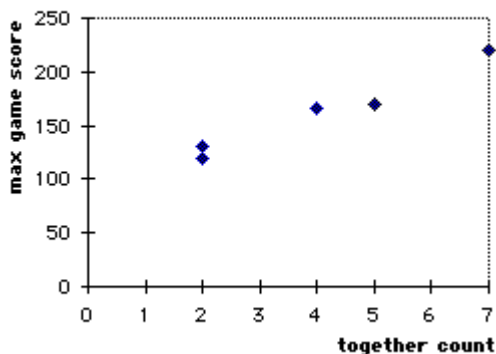

Figure 2. The Y axis shows the maximum score achieved by a player during the trials. The score is based on the number and proportion of contact times between the ring and the wire. The X axis shows the number of high (6,7) scores amongst the seven 'togetherness' questions.

Togetherness

The response variable is the number of high (6 or 7) scores out of the seven 'togetherness' questions. This is treated as a binomial variable (number of 'successes' out of 7 trials), and therefore logistic regression is used to test the influence of a linear model involving the other independent and explanatory variables on the response. All significance tests are carried out in the context of the logistic regression, and no results are reported with significance less than 5%.

For the within subjects design the significant variables are the main condition (whether haptic plus visual or just visual), gender, age, and the subjects' assessments of the social anxiety of their remote partner. The overall chi-squared for goodness of fit of the model (which should be small for a good fit) is 7.5 on 5 d.f. No variable can be deleted from the model without significantly worsening the goodness of fit.

For the between subjects design the results are the same, except that age is not significant.

To summarise, the results from the overall logistic regression are:-

- The haptic plus visual condition results in a higher sense of reported togetherness than the visual only condition.
- Females tend to report a higher sense of togetherness than males.
- Togetherness decreases with age (in the within subjects design only).
- Togetherness is positively associated with the estimated extent of social anxiety of the remote partner.

Task Performance

A performance score was constructed from the proportion of time that the ring was not intersecting the wire. Since there were many attempts for each experimental subject, we take the maximum score achieved as a measure of task performance. The table below shows the mean (standard deviation) of the maximum scores, and the results of t-tests comparing the means for the visual only conditions against the haptic plus visual. As would be expected performance is significantly better with the presence of the haptic feedback.

Treating the maximum score as a dependent variable in an ordinary multiple regression analysis results in an extremely good fit (squared multiple correlation being 97% for the within groups design and 88% for the between groups design). The model includes the basic condition and 'togetherness'. However, 'togetherness' is

significantly and positively associated with the score under the haptic plus visual condition only (ie, the regression slope for togetherness is not significant under the visual only conditions).

Table 1:

	Mean (SD) Score	Max	t-test for diff. in means
Haptic plus visual	161 (39.8)		
Visual only (within design)	67 (13.7)		P < 0.002
Visual only (between design)	77 (34.5)		P < 0.008

5. Conclusions

This note reports an ongoing experiment to examine the extent to which haptic communication, in addition to the usual visual feedback, influences the sense of togetherness between remote participants in a shared VE. Preliminary results suggest that haptic feedback adds significantly to the sense of togetherness, as does gender (higher for females). Interestingly, togetherness also increases with the degree of 'social anxiety' estimated for the other (unknown, remote) person by the subject. The reason for this is unknown, but it occurred in both experimental conditions.

There is also a clear influence of haptic feedback on the performance of the task, and independently, in the presence of haptic feedback, the degree of togetherness also significantly improves task performance.

The experiment remains to be completed, so the results above are certainly tentative.

Small Group Meetings

Summary

This paper describes an experiment that compares behaviour in small groups when they carry out a task in a virtual environment (VE) and then continue the same task in a similar real-world environment. The purpose of the experiment was not to examine task performance, but to compare various aspects of the social relations between the group members in the two environments. Ten groups of 3 people each, who had never met before, met first in a shared VE and carried out a task that required the identification and solution of puzzles presented on pieces of paper stuck around the walls of a room. The puzzle involved identifying that the same-numbered words across all the pieces of paper formed a riddle or 'saying'. The group continued this task for 15 minutes, and then stopped to answer a questionnaire. The group then reconvened in the real world, and continued the same task. The experiment also required one of the group members to continually monitor a particular one of the others in order to examine whether social discomfort could be generated within a VE. In each group there was one immersed person, with a head-mounted display and head-tracking, and two non-immersed people who experienced the environment on a workstation display. The results suggest that the immersed person tended to emerge as leader in the virtual group, but not in the real meeting. Group accord tended to be higher in the real meeting than in the virtual meeting. Socially conditioned responses such as embarrassment could be generated in the virtual meeting, even though the individuals were presented to one another by very simple avatars. The study also found a positive relationship between presence of being in a place, and co-presence, that is the sense of being with the other people. Accord in the group increased with presence, the performance of group, and the presence of females in the group. The study is seen as part of a much larger planned study, for which experiment was used to begin to understand the issued involved in comparing real and virtual meetings.

1. Introduction

There is substantial interest in the use of Virtual Environments (VEs) as a medium for collaboration between remote participants, and several systems and applications have been established to enable this, for example (Carlsson and Hagsand, 1993; Greenhalgh and Benford, 1995; Leigh andJohnson, 1996; Macedonia and Noll, 1997; Major, Stytz,Wells, 1997). There is also an explosion of virtual multi-user online worlds and communities, and the start of research into the social relations that emerge in such communities, surveyed recently by (Schroeder, 1997). However, there has been limited study of what happens when people actually make use of these systems (Bowers, Pycock, O'Brien, 1996). This paper describes an experiment, in fact part of a much larger planned experiment, that asks the question: What is the experience of participants when carrying out a task with others in a shared VE, and how similar and how different is that experience from working with these others on the same task in the real world?

The experiment was designed to explore the behaviour of small groups carrying out a task initially in a virtual and continuing in a real environment. Each of the 10 groups involved consisted of three people, unknown to one another beforehand. The group task, to be described fully later, consisted of solving a set of riddles. The task only involved observation and talking, and it could be solved most efficiently by group cooperation.

The focus of the study was not at all on performance, in the sense of how well the task was completed, but rather on how the social relations between the members developed in the virtual environment, and how, if at all, these carried over to their interactions in the real world. In particular, the study was concerned with the following issues:

- Does computational advantage confer social power?

One of the group participants was immersed in a virtual environment with a head-tracked head-mounted display, and the other two were not immersed but used a desktop workstation display. None of the participants had information as to the type of system the others were using. To what extent would the immersed person, given the empowerment bestowed by their computational advantage, become the leader of the virtual meeting, and to what extent would this carry over to the later real meeting?

- Is the sense of presence of being in the virtual place associated with 'co-presence' - the sense of being and acting with others in a virtual place?

This is a useful question to ask, since if presence and co-presence are associated this could be because of common factors influencing both, or because the individual sense of presence influences the chance of an emergent co-presence or vice versa. This was studied using reported presence based on post-experimental questionnaires.

- How does the sense of enjoyment and feelings of group affection vary as between the virtual and the real experience?

An attempt was made through questionnaire and post-experimental de-briefing to assess the extent to which the experience was 'positive', and how this changed in the transition from virtual to real.

- Can reactions such as embarrassment, shyness, conflict, be generated in the virtual environment, and if so to what extent does this carry over to the real?

In the virtual environment one of the participants was given instructions, unknown to all others, to closely follow and observe another participant. This could affect group interaction in several ways: the embarrassment of the observer, the annoyance of the observed, the sense of being left out of things by the third person.

Small group meetings in virtual environments with the people involved continuing the same task in a real environment (of which the virtual was a simulation) have not been studied before. In this experiment there was an attempt to explore the pattern of relationships within the shared VE, and also to see how these changed in continuing real meetings. The work described in this paper nevertheless makes a limited start in this endeavour - limited for two main reasons: first the length of time of the meetings was very short (15 minutes in the virtual followed by 15 minutes in the real). Second, the order in which the meetings occurred (first virtual and then real) requires a control situation where a similar number of groups carry out the experiment first in the real and then continuing in the virtual. This paper describes a study at a certain incomplete stage - nevertheless the results seemed interesting enough to report at this juncture.

The details of the experiment are given in Section 2. Results obtained by the use of post-experimental questionnaires are given in Section 3, and results from de-briefing sessions in Section 4. Section 5 discusses the results in relation to other published work, and the conlusions and way ahead are presented in Section 6.

2. Experiment

2.1 Scenario

The study involved 10 groups of three people each recruited by advertisement on the UCL campus. There was no payment for taking part in the study. The experiment took place over a two week period. There were four experimenters involved in the study, one ('minder') each to look after one of the subjects, and a 'floor manager' who maintained overall control and synchronisation of the various activities. The experiment took place in one large laboratory divided into partitions, with the three subjects at opposite sides of the laboratory. Care was taken to avoid the subjects seeing or meeting each other before the start of the experiment.

As each subject arrived they were assigned to their 'minder' who took them to their assigned workstation, or in one case to the immersive virtual reality room at one end of the main laboratory. Each subject was assigned a colour (Red, Green or Blue) and they were referred to by that colour throughout the experiment and later de-briefing. The subjects could not see their own avatars (except for the Red, immersed, person if he or she looked downwards).

Each subject was introduced to the system that they would be using. This was either a desktop system (Green and Blue) or an immersive system with a head-mounted display (Red). The virtual environment displayed was actually a rendition of the laboratory in which they were actually physically located. Each was represented by an avatar of the same colour as their assigned name.

Their first task was to individually learn to move through the environment. Then, at a signal from the 'floor manager' each subject was given a sheet describing the overall task to be performed. Then again on a signal they were invited to put on earphones, and to introduce themselves to one another. They could only refer to themselves and to the others by their colour.

The task was to locate a room which had sheets of paper stuck around the walls. The sheets each had several words in a column, each preceded by a number. The words across all sheets with a common number belonged to a 'saying' (for example, 'A critic is a man who knows the way but can't drive a car'). The task was first to figure this out and second to unscramble as many of these sayings as possible.

The subjects were asked to find the room with the papers together, and then solve the puzzle. The room with the papers was the rendition of the room with the virtual reality equipment, where the Red subject was physically located.

The Green subject was given an additional task, not revealed to the others. Green was asked to monitor Red as closely as possible, always trying to be in Red's line of vision, although taking part in the puzzle solving task as much as possible. If Red objected Green was to comply temporarily with Red's wishes, but then continue anyway with this monitoring task.

The minders sat unobtrusively near the subject throughout the virtual part of the group activity, in case of problems. The minder of Green had an additional job - to prompt Green to obstruct Red if Green did not appear to be carrying out this task but rather became only involved in the puzzle solving activity.

After about 15 minutes the virtual session was terminated, and the subjects completed a questionnaire, which took about 10 minutes. Then each subject was required to put on a waistcoat of their colour, and at a signal from the floor manager, they all met together in real life for the first time just outside the virtual reality room, the room which had the real puzzles placed on the walls.

They were then invited to continue the task in the physical location, which lasted for about another 15 minutes. At the end of that time they completed another questionnaire, and then met with the floor manager for a debriefing.

During the virtual session the virtual movements of the subjects were automatically recorded, and an audio tape recorded their conversation. The real session was videotaped from above giving a plan-view.

2.2 Materials

The Red (immersed) person was using a Silicon Graphics Onyx with twin 196 MHz R10000, Infinite Reality Graphics and 64M main memory, running Irix 6.2. The tracking system has two Polhemus Fastraks, one for the HMD and another for a 5 button 3D mouse. The helmet was a Virtual Research VR4 which has a resolution of 742×230 pixels for each eye, 170,660 colour elements and a field-of-view 67 degrees diagonal at 85% overlap.

The total scene consisted of about 3500 polygons which ran at a frame rate of no less than 20 Hz in stereo. The latency was approximately 120 ms.

The Red subject moved through the environment in gaze direction at constant velocity by pressing a thumb button on the 3D mouse. There was a virtual body (avatar) which responded to hand and head movements.

The Green subject used a SGI High Impact system with 200Mhz R4400 and 64MB main memory. The scene was shown on the full 21 inch screen display. Navigation was accomplished by using the keyboard arrow keys, with up and down arrows giving forward and back movement, and left and right keys providing rotation. All movement was on the horizontal plane of the floor.

The Blue subject used an SGI O2 running at 180Mz on Iriz 6.3, with an R5000 processor, and 32MB main memory. The scene was shown on a full 17 inch screen display. Navigation was the same as for the SGI Impact.

The sound system used was the Robust-Audio Tool (RAT) v.3.023. This allows multiple users to talk over the Mbone (Hardman, et. al., 1995).

The virtual reality software used throughout was DIVE 3.2 (Carlsson and Hagsand, 1993). A DIVE avatar was used for each of the participants, and was the same for each except for the colour. An image of such an avatar is shown in Figure 1.

Figure 1 about here.

3. Questionnaire Results

3.1 Leadership

There were two questions that related to leadership, one directly and the other indirectly. Each subject was asked to score all three subjects on the degree to which that person "was the 'leader' or main organiser" in the meeting that had just concluded. The three scores, one for Red, Green and Blue had to add to 100. In addition, there was a similar question concerning who did most of the talking.

Table 1
Mean and Standard Deviation of 'Leadership' Scores.
The 'Frequency' refers to the number of times out of 30 where

the individual had the highest leadership score.

Person	Score in Virtual	Frequency	Score in Real	Frequency
Red	46 ± 17	14	33 ± 12	5
Green	34 ± 13	2	35 ± 11	7
Blue	20 ± 13	5	32 ± 10	2

Table 1 shows the results for leadership, after the meeting in the virtual setting, and then after the real setting. The most striking aspect of this is the highly significant difference in leadership rating for Red (the immersed person) compared between the virtual and real. After the real meeting each participant was assigned approximately the same leadership rating, whereas immediately after the virtual meeting Red emerged as the clear leader. In fact 14 out of the 30 participants rated Red as the leader immediately after the virtual session, whereas 5 rated Red as leader after the real session.

Table 2

Mean and Standard Deviation of 'Most Talking' Scores
The 'Frequency' refers to the number of times out of 30 where
the individual had the highest talking score.

Person	Score in Virtual	Frequency	Score in Real	Frequency
Red	45 ± 17	16	32 ± 12	3
Green	22 ± 11	0	35 ± 11	6
Blue	33 ± 13	6	34 ± 10	5

Table 2 shows similar results for 'who did the most talking'. It is clear that Red was perceived to be the most talkative during the virtual session, but that this did not carry over to the real session. 16 of the 30 participants reported Red as the most talkative after the virtual session compared with 3 of the 30 after the Real session.

Two factors distinguish Red from Green and Blue during the virtual session. The first was that Red was 'monitored' by Green. As will be seen later, for the most part Red was unaware of this, and there is no obvious way that this could have had an effect on leadership behaviour displayed by Red. The second difference is that Red was the only one immersed through a head tracked HMD, and a hand tracker. Moreover, Red was on a machine with a faster processor. There is no way in this experiment to disambiguate the impact of the processing speed and the fact of immersion. However, Blue, the one with the least processing power, although scoring least on leadership, had the same level of talkativeness in the virtual and real experiences. The zero score for Green on talkativeness in the virtual part of the experiment probably reflects Green's additional monitoring task.

The first and perhaps most important hypothesis generated from this study is that greater computational resources may enhance leadership capability. The reported leadership behaviour of the person who was immersed vanished when all subjects participated on relatively equal terms in the real setting.

3.2 Presence and co-presence

The term 'presence' in the virtual environment literature has come to be used to denote the sense of 'being there' in a place (for example, Held and Durlach, 1992). An orthogonal attribute of presence-in-a-place, is the sense of being present with other people. This attribute is logically orthogonal, since, for example, talking on a telephone with someone might give a strong sense of 'being with them' but not of being in the same place as them. It is useful nevertheless to examine the extent to which these two different types of presence, place-presence, and co-presence, are empirically related. If they are in fact related, then this is either because they influence one another, or because there are underlying common factors to both.

The questionnaire asked the followingthree questions relating to co-presence:

1. In the last meeting, to what extent did you have the sense of the *other two people being together with you?*
2. Continue to think back about the last meeting. To what extent can you imagine yourself *being now with the other two people* in that room?
3. Please rate how closely your sense of being together with others in a real-world setting resembles your sense of being with them in the virtual room.

The following two questions related to place-presence:

4. To what extent did you have the sense of being in that room which has the pieces of paper with the riddles on the walls? (For example if you were asked this question about the room you are in now, you would give a score of 7. However, if you were asked this question about whether you were sitting in a room at home now, you would give a score of 1).
5. Think back now about the meeting and the spatial layout of the room. For example, to what extent in your imagination can you move around that room now?

Each question was rated on a 1 to 7 scale, where 1 had the legend 'Not at all' and 7 the legend 'Very much so'.

As a conservative measure of the subjective (reported) level of place- and co-presence the high scores only were taken into account. The overall measure of place-presence is the number of scores of '6' or '7', and hence is a count of 0, 1 or 2. Similarly, the overall measure of co-presence is the number of scores of '6' or '7', and hence is a count of 0, 1, 2, or 3. This approach is the same as has been used in previous studies of presence (Slater and Wilbur, 1997).

The correlation between these two scores ($r = 0.59$) is significant ($P=0.0006$). Considering only the raw scores for the two basic questions (co-presence 1 and place-presence 1) $r = 0.52$, at a similar level of significance. It is

interesting to note that the immersed person (Red) did not report a significantly higher level of presence on any category.

The second hypothesis generated from this study is therefore that presence and co-presence are linearly associated, but that the immersed person did not report a higher level of either type of presence than the other two.

3.3 Group Accord

There were several questions that attempted to assess the group members' appraisals of one another and the group as a whole. All but one question was rated on a 1 to 7 scale, where 1 meant lowest level of the quality concerned (e.g., enjoyment) and 7 meant the highest quality. In each case the overall group means and standard deviations are given for responses after the virtual and after the real setting.

For the next three questions the group scores were computed by averaging the scores over the 3 members. The significance levels are for paired t-tests over the 10 groups.

- *(enjoy)* Think about a previous time when you *enjoyed* working together in a group. To what extent have you *enjoyed* the group experience just now?

enjoyment (P = 0.003)

After Virtual	After Real
4.23 ± 1.19	5.70 ±1

- *(meet again)* Sometimes you meet people in a small group situation, and you'd like to meet them again. To what extent is the current situation similar to that?

meet again (P = 0.3)

After Virtual	After Real
4.23 ± 0.74	4.73 ±0.41

In the next case, the maximum degree of isolation was taken as the score for the group as a whole.

- *(isolated)* To what extent was anyone (including yourself) 'isolated' compared to the other two people? Give a score for each individual out of 100, where a person scores 100 if they were completely isolated from the other two, and where the three scores add to 100.

isolation (P = 0.003)

After Virtual	After Real
71.50 ± 14.54	44.40 ± 16.08

The following questions required a response by each subject for each of the other two subjects (e.g., Red would give responses with respect to Green and Blue). The score for the group is taken as the sum of the 6 scores for the individual members (six because each individual does not self-score), divided by the total possible score for the group, which is 42.

- *(individual meet again)* Would you like to meet any of the other two people again? (please put one tick in each column).

<div align="center">

(1) I would not like to meet this person

(4) No preference either way.

(7) I would very much like to meet this person

meet individuals again (P = 0.16)

</div>

After Virtual	After Real
0.66 ± 0.12	0.71 ± 0.09

- *(comfortable)* The extent to which I felt comfortable with each of the other two persons was (please put one tick in each column):

<div align="center">

(1) I felt very uncomfortable with him/her.

(4) Neither comfortable/nor uncomfortable

(7) I felt very comfortable with him/her.

comfort with others (P = 0.002)

</div>

After Virtual	After Real
0.66 ± 0.09	0.81 ± 0.1

- *(cooperative)* Overall, how cooperative were each of the other two people in the task?

<div align="center">

(1) S/he was not cooperative at all

(7) S/he was very cooperative

cooperation (P = 0.01)

</div>

After Virtual	After Real
0.77 ± 0.13	0.88 ± 0.12

- *(embarrassment)* Did any of the other two people make you feel self-conscious or embarrassed?

(1) S/he did not make me feel this way.

(7) S/he did make me feel this way very much.

embarrassment (P = 0.11)

After Virtual	After Real
0.25 ± 0.09	0.20 ± 0.06

Finally, each of the seven variables above are combined into one overall score for group 'accord'. In order to make each of the variables result in greater accord in a range from 0 to 1, the scores out of 7 are normalised to be between 0 and 1, non-isolation is taken as , and non-embarrassment is embarrassment subtracted from 1. This gives the following result:

overall accord (P = 0.000)

After Virtual	After Real
0.62 ± 0.06	0.87 ± 0.06

Taking overall group scores there is a significant difference between the result after the virtual session and after the real session, with overall group 'accord' higher after the latter. In particular, after the real session there was greater enjoyment, less isolation of individual memhers, a greater sense of comfort with the other members, and more cooperation.

The reason for the differences might not be solely due to the nature of a virtual compared to a real encounter. Another factor that was different between the two sessions was that in the virtual session Green was asked to 'monitor' Red, while this was not the case in the real session. However, when the responses for the individuals are examined, there are no significant differences between Red, Green and Blue for any of the 'accord' variables considered above.

There is also simply the question of time: after the real session the group members had been working on the puzzle altogether for about 30 minutes, compared to 15 minutes after the virtual session. This study should be considered as the first part of a larger experiment - where another 10 groups repeat the experiment but with the order of session reversed - real first and then virtual. From this study it would be possible to see if there was a significant increase in 'accord' after the second session. If so, then the result would be most likely due to time.

3.4 Accord and Presence

A previous study (Barfield and Weghorst, 1993) has found a significant relationship between presence and enjoyment. In order to examine this in relation to the current experiment a measure of individual accord was constructed on the same lines as in the previous section, except now for each individual rather than for the group as a whole. This was used as the response variable in a regression analysis where the major explanatory variables were presence, co-presence and combination of the two.

Figure 2
Accord against overall presence count

Figure 2 shows a plot of individual accord against the combined count of presence and co-presence (r=0.72). Using the combined presence count as an explanatory variable in a regression analysis, results in a significant fit, and also gender and the number of riddles solved are significant explanatory variables. Females tend to show higher accord scores than males and the more riddles solved the greater the accord.

Table 3
Multiple Regression for Accord
$R^2 = 0.66$, $t_{0.05}(28) = 2.048$

Parameter	Estimate	S.E.	t-value
Constant	0.534	0.023	
Increment in constant for females	0.086	0.034	2.230
riddles solved	0.015	0.007	2.062
Overall presence	0.056	0.009	6.008

The co-presence aspect of overall presence dominates the relationship. If co-presence only is used as the explanatory variable then a very similar result to Table 3 emerges (with $R^2 = 0.61$). If place-presence only is used, then the number of riddles is no longer significant, although gender remains so, with $R^2 = 0.45$).

3.5 Analysis of Free Responses

After the virtual session the questionnaire included the following:

- List any things that hindered you from successfully accomplishing the task.

The report concentrates only on issues that were raised by several people, rather than the more ideosyncratic comments particular to only one person. There were three common themes that were mentioned by several people that hindered them in the task: poor navigational ability, poor audio, and the discomfort of the immersed group.

Poor Navigation:

This was recorded as a problem by 8 non-immersed and 7 immersed people. The problem of 'going through walls' was especially mentioned as part of this issue.

Poor Audio:

This was mentioned by 10 people. Particular issues mentioned were delays in audio, lack of communication in the sense of being difficult to know if someone was talking to someone else, not being able to hear their own voice at normal level - tending to be too loud, not being able to easily realise who specifically was talking.

Immersion Discomfort:

This was reported by 5 of the 10 immersed people. Particular comments were: headaches, slightly out of focus, felt sick and sweaty, and 'it was physically uncomfortable experience - by the end of session was very much distracting me from the task'.

3.6 Summary

This section has examined the results of the questionnaire data. Salient hypotheses that may be generated from this study are:

- Immersion enhances leadership capability: the immersed person was overwhelmingly recognised as leader in the virtual session, but this disappeared in the real session. This was confirmed by a separate question on which person did the most talking.
- Presence (being in a place) and co-presence (being with other people) were positively correlated.
- Reported presence was not significantly different between the immersed and non-immersed people.
- Group accord increased in the real session compared to the virtual (though it is not possible in this study to rule out the effect of time).
- Higher individual accord was associated with higher overall, place- and co-presence.
- Individual accord tended to be higher for females than for males, and was positively associated with more successful performance of the task.
- There was no reported effect of the attempt to deliberately introduce some 'embarrassment' into the virtual session by having one subject monitor another - no differences between the three role-colours were reported on any component of accord.

4. Results of the Debriefing Sessions

4.1 Impact of the Monitoring Task

Questionnaires are able to capture rather static limited information about events. Often it is useful to use face-to-face unstructured encounters in order to look behind the questionnaire data and get a better understanding of what was happening - to allow for possibilities not envisaged during the questionnaire design, and to explore the dynamics of the situation. Therefore, at the end of the experiment the participants were invited for a de-briefing session, to allow them to talk freely about their experiences. In each such de-briefing the first issue for discussion was whether Red noticed anything unusual in the behaviour of the other two participants, and then the extent to which Green had found the 'monitoring' task awkward or embarrassing.

In three of the ten groups there was an impact of this additional task by Green. In Group 3 Red formed the opinion that Green was being deliberately destructive. Also in this particular session the sound from Red was 'crackling' and Green thought that Red was doing this deliberately. All three members of this group (Red and Green male, Blue female) had a high sense of what they described as 'paranoia' during the virtual session, and agreed that this completely disappeared when they met for real. This group actually never figured out even what the puzzle was, and found this to be frustrating.

In group 9 (Red and Green female, Blue male) Red did notice something different - but interpreted this as something being wrong with the avatar configurations. She said that 'Everyone was supposed to be looking at the walls, but Green was looking at me'. In this same group, Green reported that 'I felt I wasn't being me' and 'What on earth were they thinking of me?' - and found it especially difficult because she was supposed to be doing two tasks at the same time (monitoring Red and helping with the puzzle). She imagined that the other two were 'wondering why I am doing this'. Sometimes she wondered if Red would think that she were staring at her.

In group 10 (Red and Green male, Blue female), Red did not notice that Green was observing him, but did notice that the way ahead seemed to be frequently blocked. Green was not embarrassed to carry out this task. However, in this group the major impact was on Blue, who thought that Red and Green 'know where they are - I felt excluded'. In other words Blue noticed that Red and Green seemed to be close to one another most of the time, and Blue was left out of this.

One thing reported by almost all Green subjects was the difficulty of carrying out the monitoring task at all. Red moved faster than the other two subjects (on the more powerful machine and immersed). Also it was difficult for Green to know Red's field of view. There being no virtual equivalent of 'eye contact' in any meaningful sense, Green could never know whether or not Red was aware of Green's activities - there could be no 'exchange of glances'. More generally this lack of feedback about body movements and body language from the avatars was mentioned by several people.

4.2 Relationship to Avatars

A second major issue explored in the de-briefings was the relationship of the people to their avatars. The most interesting way in which this was realised was through projection - that is, individuals were respectful of the avatars of the other people, and tried to avoid carrying out actions that would cause distress or be impossible in real life.

- In Group 1 Blue said that walking through the avatar of another (which happened frequently by accident in the confined virtual space) led to his embarrassment. In the same group Red reported that walking through another body was 'weird', although Red experienced the situation as like being in 'fancy dress', the others were 'not quite real people, without a human presence, just pixels'.

- In Group 2, Green said that it 'didn't bother me to walk through people - this was the rule of this universe'. In the same group Blue found it 'frightening' to walk through a person.

- In Group 4 neither Red nor Green minded about this issue, but Blue had the impression that it was 'rude' to walk through someone.

- In Group 5 Green found it annoying if someone went through him, and Blue also thought that such it was 'bad if someone walks through you.'

- In Group 7 Red and Green each reported saying 'Sorry' when walking through through someone.

- In Group 9 Red felt it was 'disconcerting when bodies passed through each other'. Also it was 'irritating' when she 'walked back through someone and didn't know'. In the same group Green reported that she 'didn't mind going through things'. Blue said that when Red came up close to him he felt 'really uncomfortable, bloody uncomfortable', and backed off.

- In Group 10 Red 'felt like apologising' when he went through someone.

Some groups also discussed the impact of the ability to go through walls (there was no collision detection at all). In Group 1 Red felt himself to be 'panicking' when he seemed to be 'stuck in the wall'. In Group 2, Green reported that it was 'frightening' and if he did so and was outside of the scenario then this induced an 'agoraphobic' feeling. He also did not like the fact that he could not look up or down, but only straight ahead (not being immersed, there was no option to swivel the gaze direction up and down). In Group 6 Blue did not like the ability to go through walls (which was easily done by accident). Green reported the same in Group 9.

This process of being mindful of the avatars of others was surprising, they were taken seriously in spite of all their shortcomings. This relationship to the avatars was noticed in another way - the surprise that some people experienced on meeting the real person.

It was interesting to note that some of the group 'reunions' - the moment when they met for real for the first time - can only be described, unscientifically, as somewhat 'emotional'. In Group 6, Green reported a 'shock' when she really met the others. In Group 9 Red was surprised to see what Green looked like for real, and Green was

similarly surprised by the appearance of Red. In the same group Blue found surprising the shape of the others' heads - somehow he had expected these to be the same as in the virtual session!

4.3 Summary

This analysis of the post-experimental group discussion revealed a surprising degree of attachment and relationship towards the virtual bodies (avatars). Although, except by inference, the individuals were not aware of the appearance of their own body, they seemed to generally respect the avatars of others, trying to avoid passing through them, and sometimes apologising when they did so. These were very simple avatars, with limited movement and no capability for any kind of emotional expression. If even these can evoke such reponses, it is interesting to wonder what responses more powerful avatar representations might evoke.

5. Discussion

5.1 Why Shared VEs Are Needed

The need for shared VEs for collaborative working is not obvious - clearly multimedia systems with real-time video and audio are capable of bringing remote people together for collaborative work. It could be argued that such multimedia systems are not suitable where there is a requirement for manipulation of objects, or shared design - although whiteboards go a long way in helping with such tasks. A study is considered in this section where even though the task does not involve shared design or manipulation, the results strongly suggest that a shared VE might offer substantial benefits. Isaacs et. al. (1995) describe an experiment using the Forum system, which compares face-to-face with distributed presentations. The application involved people giving presentations to groups. There were 14 presentations, half given by the presenter in a lecture hall with the audience in conventional style, and half given using the Forum system, a desktop based video and audio system. The presentations were paired so that the presenter gave the same material twice, once to an audience in a face-to-face lecture hall setting, and the other to a different distributed audience using the Forum system.

The Forum involved live video, audio and slides presented on a desktop workstation. The audience members could see live video of the presenter, and the slides (which could be followed along with the speaker, independently scrolled and annotated by the audience members). The audience members could speak to the whole group, and send messages to the speaker and one another. The speaker could not see the audience, and the audience members could not see one another.

The Forum audiences could be using other applications on their desktop machine during the session, whereas of course the lecture attendees had to physically travel to the meeting place, and could not easily be engaged in other activities during the course of the lecture. Hence the Forum audiences tended to be larger, and also self-selected.

The important results from the point of view of this paper concerned the perceived quality of the presentation both from the point of view of the presenter and the audience. The Forum talks tended to be longer than the real talks, because the speakers lost track of time. The speakers reported that they were unable to see the usual audience cues of increased restlessness at around the time the talks were scheduled to complete. Generally, the speakers had a weak sense of audience reaction, since they were unable to see or hear the usual types of subtle audience responses in the course of a lecture. The experimentors noted that sometimes during the Forum talks audience members did spontaneously chuckle and applause, but of course neither the speaker nor other audience members were aware of this. Overall the Forum did not provide sufficient support for the cues that speakers use to monitor and adjust to audiences.

During the course of a face-to-face lecture, a speaker might call on an audience member to help in discussion of a particular point, especially where that audience member was known to have special knowledge of the particular issue. In the case of the Forum, the speaker was reluctant to ask someone in the displayed audience list to contribute in this way, because there was no guarantee that the person was paying attention - they could at that moment have been using some other application on their workstation, and there was no indication of this. More generally, the speakers complained that they did not get the immediate feedback they usually rely upon when answering a particular question for someone, such as seeing them nod or shake their head, or the expression on their face.

The essential point is that although the audience and speakers are together in a shared system, the space that they inhabit together is fragmented between a video representation of the speaker, the audio channel, the lists of audience members, and the workstation environment. There is no unified common space with a metric where participants can vary the distances between one another and become aware of changing spatial relationships, and responses to those changes. In particular, although there is visual representation, there is no visual space which all participants simultaneously inhabit. There are no dynamic representations of individuals (except for the presenter) to which other individuals can relate and respond, and know that their responses may be experienced by others.

In spite of the current technical shortcomings, shared VEs do offer a common shared visual space, an ideally synchronised audio space, ideally a common haptic space, with ideally multi-modal (vision, audio, haptic) personal representations - the 'avatars'.

5.2 Some Characteristics of Shared VEs

The idea of a unified shared space and avatar representations in a shared VE is supported by McQuaid in Nunamaker (1997), in the context of group support systems. In particular he agues that avatars can give other participants a way to judge the focus of attention of others, for example, seeing when two other people are directly communicating. He suggests that avatars can convey information that is given by physical movement in the real world, and that in VEs avatar configurations may take on different social meaning than in everyday reality. For example, sitting on a chair in real life is for comfort and relaxation, and to facilitate certain types of activity. In VEs there is no inherent need for avatars not directly mapped to the actions of their real human counterparts to

sit. Yet 'sitting' might take on the meaning that the real human counterpart of a seated avatar is currently otherwise engaged and not actually present in the VE. Of course avatars can also exhibit movements that have a social meaning directly mapped from everyday reality - in the context of a VE lecture, avatars can be made to nod, shake their head, exhibit facial expressions, become fidgety, giving cues to the speaker about audience reactions. The experiment which is the subject of this paper suggests that even where there are very simple non-expressive avatars, that social conventions may carry over - people can become embarrassed or angry while embodied in very basic avatars, and treat each other's avatars with care. This is a necessary (but not sufficient) condition for social interaction and group working within a shared VE.

5.3 Some Characteristics of Avatars in Shared VEs

There are two characteristics in the experimental setup described in this paper that can easily be overlooked, but are actually worth questioning. The first is that the experiment was carried out in a virtual copy of a real laboratory environment, i.e., a virtual reflection of a real spatial organisation. The second characteristic, is that the participants were represented by avatars that had a humanoid resemblance at least, if minimal human body functionality. Given the nature of virtual environments, neither of these characteristics are necessary - there is no need to organise virtual space to be anything like real space, and no intrinsic need for participants to be virtually embodied, or embodied with a human appearance. Yet these are characteristics generally employed in shared virtual environments.

Given that there is a common space that is inhabited by avatars, what characteristics and capabilities should these have? Rich et. al. (1994) describe a shared VE system for "learning by doing" a world which it is possible to explore and learn to use athletic equipment, and configured as an on-line community.

There is a virtual body controlled by a user, and also an artificial agent also embodied as an avatar. Generally agents (the humans, virtual humans and other virtual beings such as birds) are able to generate sound, and move as expressive articulated figures. The human avatars had independently controllable head, torso and forearms controlled via an actuator system. The goal was to make the users feel as if they were inhabiting a body rather than just operating an animated figure. It was argued that this was achieved by the ability of users to control navigation through hand gestures based on a video recognition system, and posture, the changing configuration of different body parts, through a switch box and joy stick. No experimental evidence of the outcome was reported.

Benford et. al. (1994) discuss extensively the social significance of space as a resource for activity and interaction in VEs. In fact much of what they say actually is to do with the activity of avatars in space, rather than just with space in itself. They argue that continual awareness of others allows people to flexibly modify their own behaviour in social situations - for example, someone heading across the room towards another probably indicates an interest in starting up a verbal communication. They describe how the use of space, or rather the avatars in a meaningful spatial configuration, allows the support or indeed emergence of social mechanisms for control of scarce resources. In a public debate a 'line' can form around a podium showing to everyone which and how many people are preparing to speak, who indeed the current speaker is (floor control), and the audience reactions (for example, they could all 'walk out') to an uninteresting talk (something that would be clearly noticed by the

speaker, unlike in the Forum system). The authors describe in detail mechanisms that can be provided by the VE to facilitate social interaction above and beyond just copying basic real-life mechanisms, in their notions of aura (the bounding presence of an object), focus (the field in which a user can become aware of others), and nimbus (the field in which others can become aware of the user). They show that social interactions can be seen as a form of negotiation between agents based on their aura, focus and nimbus fields.

In their discussion Benford et. al. again emphase the importance of embodiment - how this can provide information about the identity and activity of the participant, how gesture and facial expression can be used for the expression of emotion, and the separation of 'mind' and 'body' - that is how the avatar can be used to signal that the real person is currently no longer 'present' in the VE but engaged in other activities (e.g., by presenting a 'sleeping' avatar).

In a later paper by Bowers et. al. (1996) there is an empirical study of what actually happens in an unstructured small group virtual meeting based on the MASSIVE system (Greenhalg and Benford, 1995). The emphasis was on understanding the relationship between the embodiment of participants through their 'blocky' avatars, and communication issues such as turn taking while talking, and other aspects of social interaction. The study used Conversational Analysis to transcribe conversation and was extended to include the simple avatar 'gestures' possible in the system (such as whole body turning or 'ear flapping'). The study found that in spite of the very limited repertoire of the avatars, the avatars were nevertheless sometimes used to supplement language as an additional mechanism in social interaction. The avatars were not just a means of navigation and representation, but became invested with social meaning, a finding that supports the results of the experiment described in this paper.

The Bowers study also found that particpants did move their avatars in socially meaningul ways, for example, to get a better view of those with whom they wished to interact. Partipants sought 'face-to-face' communication, even though the use of the audio channel did not actually require this. Although talk was accompanied by the limited repertoire of gestures only to a very limited extent, they did find that there was mutually coordinated movements amongst two or more participants. This suggested that embodiments should support higher order activities than mere movement, actions of social significance, such as approaches, exchange of glance, turning to, turning away, and other basic expressive actions.

The latter requirements are fully supported by the current study - recall that, for example, the Green subject found it difficult to know whether their monitoring task was effective because it was hard to tell whether or not they had been noticed by Red. Even more fundamental - it was hard for participants to tell which subject was talking, because there were no accompanying lip movements, and no spatialised sound. On this latter point Rich et. al. argued that crude images together with crude audio rendering provides better feedback to participants than better visual or better audio by themselves. They give an example from their system of the avatar walking. Without shadows it is impossible to tell if the avatar is actually walking along on the floor, and with spatialised sound it is only possible to tell whether the walking noise is coming from the left and right. But when sound is combined with the visual rendering, the brain seemlessly integrates the two into a 'foot stepping on the floor' totality, so that the participant can tell exactly when each foot strikes the floor. In the context of avatars talking, even

crude lip movement without spatialised sound is likely to give very strong feedback about who is currently talking.

Vilhjálmsson (1997) provides an elegant approach to avatar functionality in his BodyChat system. He argues that the avatar behaviour should be encapsulated into layers, and that at the bottom layer there are what might be described as fundamental or autonomic behaviours that are always happening. This not only gives the sense of 'aliveness' of each avatar individually, but also enhances the ability of people to interact. So at the very basic level, avatars visibly breath. Avatars have large black eyes, but with a 'twinkle' in the centre. When one of these avatars 'looks' at you, there is a sense that there is some presence there. Each avatar can be in a state of 'Being Available' or not being so. When two avatars pass each other while walking, they will carry out an involuntary glance towards each other. There is no doubt that each is in the field of view of the other. What happens subsequently is an automatic negotiation based on the state of availability of each. For example, if both are available, then they may stop walking and the potentiality of a conversation ensues. During a conversation there are subtle cues - for example, raised eyebrows for questions, and not-so-subtle cues, such as corresponding lip movements. In fact many of the complaints of the subjects in the experiment of this paper, would have been overcome through the use of BodyChat, and this without any particularly complex body representations - in a computer graphics sense the BodyChat avatars are no more geometrically complex than those available in DIVE. This idea that the avatars systems themselves take care of many autonomic responses, of which in real life we are hardly aware, if aware at all, seems an excellent way forward in the design of personally and socially meaningful embodiments. There is some empirical evidence for this, in a study carried out by Thorisson and Cassell (1996) who conclude that: "This supports our claim that what really matters in face-to-face dialogue is, in addition to "classical information exchange", the supportive behaviors that often have been dismissed as incidental to effective interaction".

To conclude this section recall that the current experiment found that the avatars had social significance even with the essentially lifeless avatars that were used - how much more compelling might the experience be with the BodyChat concepts employed?

5.4 Presence, Immersion, and Quesionnaires

In other works (summarised in Slater and Wilbur, 1997) the authors have argued that a question of interest in the study of VEs is the relationship between the immersive capabilities of a system, and the behavioural and psychological implications of this in terms of presence. The underlying hypothesis is that the more immersive experience the system can deliver (a surrounding environment where energy from each sensory modality can arrive at the participant from any direction) is likely to generate a greater and measurable sense of presence - the question of interest being the balance of immersive attributes are needed to achieve particular levels of presence (Ellis, 1996). It would follow from this that the Red people in this study ought to have reported higher levels of presence than Green and Blue - but this was not the case.

Of course, this may be so because the Red people did not generally have a higher sense of presence, for whatever reason. However, another aspect of this study indicates another way to think about this. Recall that the impact of

Green's monitoring task was only uncovered through the post-experimental discussion (the de-briefing session). This could not have been discovered from the questionnaires, even though there were questions relating to the accord within the group. There were no differences in questionnaire response in accord between Red, Green and Blue. As stated earlier, questionnaires are limited tools in capturing relatively static attributes of a situation - dynamics and reasons are unlikely to be uncovered. The same is likely to be the case with presence - questionnaires are useful in making a start in uncovering this phenomenon, but certainly should be supplemented by more in-depth elicitation techniques, which, even if less amenable to quantification might give much greater insight.

This is not presented as an explanation for the result, but rather as a failure of technique, which needs to be put right in subsequent studies.

6. Conclusions

This paper presents an experimental study comparing small group behaviour while carrying out a task in a shared VE, and then continuing the task in the real world, where the VE was a virtual copy of the real-world environment. The results of the study, bound as they are by the specific conditions of this experiment, suggest the following hypotheses for future research:

- Leadership capability is enhanced by computational power - in particular it may be that leadership is enhanced by greater levels of immersion.

- Personal responses to social situations, such as embarrassment, discomfort, can be generated in a shared VE, even though the people involved are experiencing one another through very limited personal avatar representations.

- Even very limited avatars take on social significance, and people have a tendency to be respectful of each other's avatars.

- Presence and co-presence are positively associated, though the causality is unknown, and better techniques for eliciting these factors are required.

- Collision detection, enabling avatars to easily obey spatial boundaries (not going through walls) and avoiding one another, must be a crucial component of any shared VE that adopts a conventional spatial representation.

As has been mentioned this paper reports a partial study. More groups must be included, the contact time must be extended, the order of presentation varied (some groups should meet first in the real world and then continue in the VE), the monitoring task or its equivalent in a future study, should be included in both real and virtual parts, and for some groups only in the real part, rather than just the particular configuration used here. Essentially, this study was conducted to find out some of the questions that should be asked in a more thorough and extensive experiment, and the results should be considered in that spirit.

CHAPTER 12   *References*

TABLE 2.

Aserinksky, E., Kleitman, N. (1953) Regularly Occurring Periods of Eye Motility and Concomitant Phenomena During Sleep, Science 118, 273-74.

Baker, R.J. (1986) The GLIM System: User's Guide, Release 3.77, Royal Statistical Society, Numerical Algorithms Group (NAG).

Bandler, R. and J. Grinder (1975) The Structure of Magic I, Science and Behaviour Books.

Bandler, R. and J. Grinder (1979) Frogs into Princes, Real People Press.

Barfield, W. and Hendrix, C. (1995) The Effect of Update Rate on the Sense of Presence within Virtual Environments, Virtual Reality: The Journal of the Virtual Reality Society, 1(1) 3-16.

Barfield, W. and S. Weghorst (1993) The Sense of Presence Within Virtual Environments: A Conceptual Framework, in Human-Computer Interaction: Software and Hardware Interfaces, Vol B, edited by G. Salvendy and M. Smith, ElsevierPublisher, 699-704, 1993.

Barfield, W. Sheridan, T., Zeltzer, D. and Slater, M. (1995) Presence and Performance Within Virtual Environments, in W. Barfield and T. Furness (eds) Virtual Environments and Advanced Interface Design, Oxford University Press.

Benford, S., Bowers, J., Fahlen, L.E., Mariani, J., Rodden, T. (1994) Supporting Cooperative Work in Virtual Environments, The Computer Journal 37(8), 653-668.

Benford, S.; Greenhalgh, C.; Bowers, J.; Snowdon, D.; Fahlen, L.E. (1995) User embodiment in collaborative virtual environments, Human Factors in Computing Systems. CHI'95 Conference, 242-9.

TABLE 2.

Bishop, G. et. al. (1992) Research Directions in Virtual Environments: Report of an NSF Invitational Workshop, Computer Graphics, 26(3), 153-177.

Bly, S., Harrison, S. and Irwin, S. (1993) Media Spaces: brining people together in a video, audio and computing environment. Communications of the ACM, 36(1), 28-47.

Bowers, J., Pycock, J., O'Brien, J. (1996) Talk and Embodiment in Collaborative Virtual Environments, CHI'96 Electronic Proceedings http://www.acm.org/sigchi/chi96/proceedings/papers/Bowers/jb_txt.htm

Brooks, F.P. et. al. (1992) Final Technical Report: Walkthrough Project, Six Generations of Building Walk-throughs, Department of Computer Science, University of North Carolina, Chapel Hill, N.C. 27599-3175.

Brooks, F.P. Jr, Ouh-Young Ming and J. Batter (1990) Project Grope - haptic displays for scientific visualization, Computer Graphics 24(4) 177-185.

Buckner, M., N.M. Meara, E.J. Reese, M. Reese (1987) Eye movement as an indicator of sensory components in thought, Journal of Counseling Psychology, 34(3) 283-287.

Carlsson, C. and Hagsand, O., DIVE A Platform for Multi-User Virtual Environment, Computer & Graphics Vol 17, No. 6, 1993, pp. 663-669.

Cavanagh, P., Leclerc, Yvan G. (1989) Shape from shadows. Journal of Experimental Psychology: Human Perception & Performance, 15 (1), 3-27.

Chin, N., Feiner, S. (1989) Near Real-Time Shadow Generation Using BSP Trees, Computer Graphics 23(3), 99-106.

Chrysanthou G. and M. Slater (1992) Dynamic Changes to Scenes Represented as BSP Trees, Eurographics 92, Computer graphics Forum, 11(3), ed. A. Kilgour and L. Kjelldahl, pp321-332.

Chrysanthou, Y., Slater, M. (1995) Shadow Volume BSP Trees for Computation of Shadows in Dynamic Scene, ACM SIGGRAPH Symposium on Interactive 3D Graphics (April, 1995).

Chung, J. (1992) A Comparison of Head-tracked and Non-head-tracked Steering Modes in the Targeting of Radio-therapy Treatment Beams, Proceedings of the 1992 ACM Symposium on Interactive 3D Graphics, 193-196.

Cohen, K.M. (1982) Speaker interaction: Video teleconferences versus face-to-face meetings. Proceedings of Teleconferencing and Electronic Communications, University of Wisconsin, 189-199.

Cool, C., Fish R.S., Kraut R.E. and Lowery C.M. (1992) Iterative Design of Video Communication Systems, in Proceedings of CSCW'92, Toronto, Canada.

Cox, D.R. (1970) Analysis of Binary Data, London: Menthuen.

DeLozier, J. and J. Grinder (1989) Turtles All The Way Down, Meta Publications.

DEVRL group (1995) Distributed Extensible Virtual Reality Laboratory (DEVRL):A Project for Cooperation in Multi-Participant Environments, submitted. (The DEVRL group consists of Departments of Computer Science at The Universities of Lancaster and Nottingham, University College London, and QMW. The Principal Investigators are S. Benford, T. Rodden, M. Slater, S. Wilbur.

TABLE 2.

Dilts, R., Grinder,J., Bandler, R., DeLozier, J., Cameran-Bandler, L. (1979) Neuro-Linguistic Programming I, Meta Publications.

Durlach, N., Slater, M., 1998, "Presence in shared virtual environments and virtual togetherness", Workshop on Presence in Shared Virtual Environments, BT Laboratories, June, 1998, http://www.cs.ucl.ac.uk/staff/m.slater/BTWorkshop/

Einspruch, E.L. and B.D. Forman (1985) Observations concerning research literature on neuro-linguistic programming, Journal of Counseling Psychology, 32(4) 589-596.

Ellis, S.R. (1991) Nature and Origin of Virtual Environments: A Bibliographic Essay, Computing Systems in Engineering, 2(4), 321-347.

Ellis, S.R. (1996) Presence of Mind: A Reaction to Thomas Sheridan's "Further Musings on the Psychophysics of Presence", Presence: Teleoperators and Virtual Environments, MIT Press, 5(2), 247-259.

Fahlen, L. (1992). The MultiG Telepresence System, Proceedings of the 4th Multi-G Workshop, Stockholm-Kista, May 1992.

Fahlen, L. (1993) Virtual Reality and the MultG Project, Virtual Reality International 93, Proceedings of the Third Annual Conference on Virtual Reality, Meckler, London, 78-86.

Fairchild, K.M., Beng Hai Lee, J. Loo, H. Ng, L. Serra (1993) The Heaven and Earth Virtual Reality: Designing Applications for Novice Users, IEEE Virtual Reality Annual International Symposium (VRAIS), September 18-22, Seattle, Washington, 47-53.

Fisher S. (1986) Telepresence master glove controller for dexterous robotic end-effectors, SPIE Intelligent Robots and Computer Vision.

Fisher, S. (1982) Viewpoint Dependent Imaging: An Interactive Stereoscopic Display, Proceedings SPIE 367, 41-45.

Fisher, S., M. McGreevy, J. Humphries, W. Robinett (1986) Virtual Environment Display System, ACM 1986 Workshop on 3D Interactive Graphics, Chapel Hill, North Carolina, October 23-24.

Foley, J.D. (1987) Interfaces For Advanced Computing, Scientific American, 257(4), October, 126-135.

Gale, S. (1990) Human aspects of interactive multimedia communication. Interacting with Computers, 2. 175-189.

Gibson, E.J. and R.D. Walk (1960) The "visual cliff", Scientific American, 202, 64-71.

Gibson, J.J. (1986) The Ecological Approach to Visual Perception, Lawrence Erlbaum Associations, Publishers, New Jersey.

Girgus, J.J., Rock, I., Egatz, R. (1977) The Effect of Knowledge of Reversibility on the Reversibility of Ambiguous Figures, Perception and Psychophysics, 22(6), 550-556.

Greenhalgh, C. and Benford, S. (1995) MASSIVE: A Collaborative Virtual Environment for Teleconferencing, ACM Transactions on Computer Human-Interaction (TOCHI), 2(53) September, 239-261.

Greenhalgh, C. M., and Benford, S. D. (1995) MASSIVE: A Virtual Reality System for Tele-conferencing, ACM Transactions on Computer Human Interfaces (TOCHI), 2 (3), 239-261.

TABLE 2.

Gregory, R.L. (1990) Eye and Brain: The Psychology of Seeing, Fourth Edition, Weidenfield and Nicholson, 182-187.

Grimsdale, C. (1991). dVS - Distributed Virtual environment System, Proceedings of Computer Graphics 91 Conference, London. ISBN 0 86353 282 9

Grinder J. and Bandler R. (1981) Trance-formations, Real People Press, Utah, ISBN 0-911226-23-0.

Grinder, J. and Bandler, R. (1976) The Structure of Magic II, Science and Behaviour Books.

Hardman, V., Sasse, M-A., Handley, M., and Watson, A. (1995) "Reliable Audio

Harrison, S. (1992) Making a Place in Media Space. Technical Report SSL-92-42. Xerox PARC, Palo Alto, CA.

Heeter, C. (1992) Being There: The Subjective Experience of Presence, Presence: Teleoperators and Virtual Environments, 1(2), MIT Press, 262-271.

Heeter, C. (1992) Being There: The Subjective Experience of Presence, Telepresence, Presence: Teleoperators and Virtual Environments, 1(2), spring 1992, MIT Press, 262-271.

Held, R.M. and N.I. Durlach (1992) Telepresence, Presence: Teleoperators and Virtual Environments, 1, winter 1992, MIT Press, 109-112.

Hendrix, C. and Barfield, W. (1996a) Presence within Virtual Environments as a Function of Visual Display Parameters, Presence: Teleoperators and Virtual Environments, MIT Press, 5(3), 274-289.

Hendrix, C. and Barfield, W. (1996b) The Sense of Presence within Auditory Environments, Presence: Teleoperators and Virtual Environments, MIT Press, 5(3), 290-301.

Hertz, J., A. Krogh, R.G. Palmer (1991) Introduction to the Theory of Neural Computation, Addison-Wesley Publishing Company.

Ho, C., Basdogan, C., Srinivasan, M.A., 1997, "Haptic Rendering: Point - and Ray - Based Interactions", *Proceedings of the Second PHANToM Users Group Workshop*, Dedham, MA, Oct. 20-21.

Hoaglin, D. and Welsch, R. (1978) The hat matrix in regression and ANOVA, American Statistician, 32, 17-22, 132.

Hodges, L.F., Rothbaum, B.O., Watson, B., Kessler, G.D., Opdyke, D. (1996) A Virtual Airplane for Fear of Flying Therapy, IEEE 1996 Virtual Reality Annual International Symposium, 86-94.

Hodges, L.F., Rothbaum, B.O., Watson, B.A., Kessler, G.A., Opdyke, D. (1995) A Virtual Airplane for Fear of Flying Therapy, VRAIS'95 [full reference to be supplied!]

Isaacs, E. A., Tang, J. C., (1993) What video can and can't do for collaboration: a case study. Proceedings of ACM Multimedia '93, Anaheim, Ca. 199-207.

Isaacs, E.A.; Morris, T.; Rodriguez, T.K.; Tang, J.C. (1995) A comparison of face-to-face and distributed presentations, Human Factors in Computing Systems. CHI'95 Conference Proceedings, 354-61.

Iwata, H. and K. Matsuda (1992) Haptic Walkthrough Simulator: Its Design and Application to Studies on Cognitive Map, The Second International Conference on Artificial Reality and Tele-existence, ICAT 92, 185-192.

TABLE 2.

Kalawsky, R. (1993) The Science of Virtual Reality and Virtual Environments: A Technical, Scientific and Engineering Reference on Virtual Environments, Addison-Wesley Publishing Company.

Karlin, S. (1969) A First Course in Stochastic Processes, Academic Press.

Kendall, M. [1975] Multivariate Analysis, Charles Griffin & Co. Ltd.

Kohler, W. (1959) Gestalt psychology: An introduction to new concepts in modern psychology. NY: Liverright, 1947; NY:NAL, 1959.

Lankton, S.R. (1979) Practical Magic: The Clinical Application of Neuro-Linguistic Programming, Meta Publications.

Leary, M., 1983, "Social Anxiousness: The construct and its measurement", Journal of Personality Assessment, Vol. 47, pp. 66-75.

Leigh, J.; Johnson, A.E. (1996) Supporting transcontinental collaborative work in persistent virtual environments, IEEE Computer Graphics and Applications, 16(4) 47-51.

Lewis S. and Mateas, M. (1994) Position paper for TelePresence workshop, ACM Multimedia '94, San Francisco, Ca. 17-19.

Linakis, V. (1995) A Study on the Transfer of Skills using a Virtual Tri-Dimensional Chess System, Advanced MSc Thesis, Department of Computer Science, QMW University of London.

Loomis, J.M. (1992a) Distal Attribution and Presence, Presence: Teleoperators and Virtual Environments, 1, winter 1992, MIT Press, 113-119.

Loomis, J.M. (1992b) Presence and Distal Attribution: Phenomenology, determinants, and assessment, SPIE 1666 Human Vision, Visual Processing and Digital Display III, 590-594.

Lusted, H.S., R.B. Knapp and A. Lloyd (1992) Biosignal Processing in Virtual Reality, Presented at the 3rd Annual Virtual Reality Conference, San Jose, CA, 25th Sept. 1992.

Lusted, H.S., R.B. Knapp and A. Lloyd (1993) Applications for Biosignal Processing in Virtual Reality, VR 93, Virtual Reality International, Proceedings of he third annual conference on Virtual Reality, London, pp134-137. Meckler.

Macedonia, M. R., Zyda, M. J., Pratt, D. R., Barham, P. T. and Zeswitz, S. (1994) NPSNET: a network software architecture for large scale virtual environments, Presence, 3(4), MIT Press.

Macedonia, M.R.; Noll, S. (1997) A transatlantic research and development environment, IEEE Computer Graphics and Applications, 17(2) 76-82.

Mackinlay, J.D., S.K. Card, G.G. Robertson (1990) Rapid Controlled Movement Through a Virtual 3D Workspace, Computer Graphics (SIGGRAPH) 24(4), 171-176.

Major, S.B.; Stytz, M.R.; Wells, W.D. (1997) Requirements and implementation of multiple modes of collaboration and cooperation within distributed virtual environments, Military, Government, and Aerospace Simulation. Proceedings of the 1997 Simulation MultiConference, Chinni, M.J. ed., Atlanta, GA, USA; 6-10 April 1997, 175-80.

TABLE 2.

McCauley, M.E., T.J. Sharkey (1993) Cybersickness: Perception of Self-motion in Virtual Environments, Presence: Teleoperators and Virtual Environments, 1(3), 311-318.

McCullagh, P. and Nelder, J.A. (1983) Generalised Linear Models, Chapman and Hall.

McGreevy, M.W. (1993) The Presence of Field Geologists in a Mars-Like Terrain, Presence: Teleoperators and Virtual Environments, 1(4) fall 1993, MIT Press, 375-403.

Mee, A. (1898) The Pleasure Telephone, the Strand Magazine, 339-369.

Mine, M.R., Brooks, F.P, Jr., Sequin, C. (1997) Moving Objects in Space: Exploiting Proprioception in Virtual-Environment Interaction, Computer Graphics (SIGGRAPH) Proceedings, 19-26.

Mizell, D., Jones, S., Jackson, P. and Picket, D. (1995) Is VR Better than a Workstation: A Report on Human Performance Experiments in Progress, Eurographics Workshop on Virtual Environments, Proceedings ed. M. Goebel, Monte Carlo, Jan 31-Feb 1st, 1995.

Monguio-Vecino, I., L.G. Lippman (1987) Image formation as related to visual fixation point, Journal of Mental Imagery, 11(1) 87-96.

Nunamaker Jr, J.F. (1997) Future research in group support systems: needs, some questions and possible directions, Int. J. Human-Computer Studies, 47, 357-385.

O'Conaill, B., Whittaker, S., and Wilbur, S. (1993) Conversations Over Video Conferences: An Evaluation of the Spoken Aspects of Video-Mediated Communication. Journal of Human-Computer Interaction, Vol. 8. 389-428.

Ochsman, R.B and Chapanis, A. (1974) The effects of IO communication modes on the behavior of teams during cooperative problem-solving. International Journal of Man-Machine Studies, 6. 579-619.

Patel, H. (1994) The Influence of Sound on Presence in Virtual Environments, MSc Advanced Methods in Computer Science, Project Report, Department of Computer Science, QMW University of London.

Pausch, R., Proffit, D., Williams, G. (1997) Quantifying Immersion in Virtual Reality, Computer Graphics (SIGGRAPH) Proceedings, 13-18.

Pausch, R., Shackelford, M.A. and Proffitt, D. (1993) IEEE Symposium on Research Frontiers in Virtual Reality, 41-45.

Pimentel, K. and K. Teixeira (1993) Virtual Reality: Through the new looking glass, McGraw-Hill.

Prothero, J., Parker, D., Furness, T. and Wells, M. (1995).Towards a Robust, Quantitative Measure for Presence.. In Proceedings of the Conference on Experimental Analysis and Measurement of Situation Awareness, 359-366.

Puerta, A.M. (1989) The power of shadows: shadow stereopsis. Journal of the Optical Society of America, Feb. 1989, 6, 309 - 311.

Rich, C., Waters, R.C., Schabes, Y., Freeman, W.T., Torrance, M.C., Golding, A.R., Roth, M. (1994) An Animated On-Line Community with Artificial Agents, IEEE Multimedia 1(4), 32-42.

Robinett, W. and R. Holloway (1992) Implementation of Flying, Scaling and Grabbing in Virtual Worlds, ACM Symposium on Interactive 3D Graphics, Cambridge MA.

TABLE 2.

Robinett, W., R. Holloway (1992) Implementation of Flying, Scaling and Grabbing in Virtual Worlds, ACM Symposium on Interactive 3D Graphics, Cambridge MA.

Rothbaum, B., Hodges, L., Kooper, R., Opdyke, D., Willford, J., North, M. (1995) Effectiveness of Computer Generated (Virtual Reality) Graded Exposure in the Treatment of Acrophobia, American Journal of Psychiatry, 152(4), 626-628.

Rutter, D.R., Stephenson, G.M. and Dewey, M.E. (1981) Visual communication and the content and style of conversations, British Journal of Social Psychology. 20, 41-52.

Sacks, H., Cchegloff, E., Jefferson G. (1974) A simplest systematics for the organisation of turn-taking in conversation, Language, 50, 696-735.

Sacks, O. (1985) The Man Who Mistook His Wife for a Hat, Picador.

Schroeder , R. (1997) 'Networked Worlds: Social Aspects of Multi-User Virtual Reality Technology' Sociological Research Online, vol. 2, no. 4, http://www.socresonline.org.uk/socresonline/2/4/5.html

Seligmann, D. D., Mercuri, R. T., and Edmark, J. T. (1995) Providing Assurances in a Multimedia Interactive Environment, in Proceedings of ACM CHI'95, Denver, Colarado.

Sharpley, C.F. (1987) Research findings on neurolinguistic programming: Nonsupportive data or an untestable theory? Journal of Counseling Psychology, 34(1), 103-107.

Sheridan, T.B. (1992) Musings on Telepresence and Virtual Presence, Presence: Teleoperators and Virtual Environments, 1, winter 1992, MIT Press, 120-126.

Sheridan, T.B. (1996) Further Musings on the Psychophysics of Presence, Presence: Teleoperators and Virtual Environments, MIT Press, 5(2), 241-246.

SIGGRAPH 89, Panel Proceedings (1989) Virtual Environments and Interactivity: Windows to the Future, Computer Graphics 23(5) 7-38.

SIGGRAPH 90, Panel Proceedings (1990) Special Session, Hip, Hype and Hope - The Three Faces of Virtual Worlds.

Slater, M. (1995) The Proceedings of the FIVE'95 Conference, Department of Computer Science, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK.

Slater, M. and Usoh, M. (1993) Presence in Immersive Virtual Environments, IEEE Virtual Reality Annual International Symposium (VRAIS), September 18-22, Seattle, Washington, 90-96.

Slater, M. and Usoh, M. (1993a) Presence in Immersive Virtual Environments, Proceedings of the IEEE Conference - Virtual Reality Annual International Symposium, IEEE Neural Networks Council, Seattle, WA (September, 1993), 90-96.

Slater, M. and Usoh, M. (1992) An Experimental Exploration of Presence in Virtual Environments, Department of Computer Science, QMW University of London.

Slater, M. and Usoh, M. (1993b) Simulating Peripheral Vision in Immersive Virtual Environments, Computers and Graphics, November 1993.

TABLE 2.

Slater, M. and Usoh, M. (1994) Representation Systems, Perceptual Position and Presence in Virtual Environments, Presence: Teleoperators and Virtual Environments, 2.3 MIT Press, 221-234.

Slater, M. and Usoh, M. (1995) Modeling in Immersive Virtual Environments: A Case for the Science of VR, in R. Earnshaw, J. Vince, H. Jones (eds) Virtual Reality Applications Academic Press, 53-70.

Slater, M., A. Steed and M. Usoh (1993) The Virtual Treadmill: A Naturalistic Metaphor for Navigation in Immersive Virtual Environments, First Eurographics Workshop on Virtual Reality, ed. by M. Goebel, 71-86.

Slater, M., Usoh, M., Steed, A. (1994) Depth of Presence in Immersive Virtual Environments, Presence: Teleoperators and Virtual Environments, MIT Press 3(2), 130-144.

Slater, M., C. Alberto, M. Usoh (1995) In the Building or Through the Window? An Experimental Comparison of Immersive and Non-Immersive Walkthroughs, Virtual Reality Environments in Architecture, Leeds 2-3rd November 1994, Computer Graphics Society.

Slater, M., Linakis, V., Usoh, M., Kooper, R. (1996) Immersion, Presence and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chesss, ACM Virtual Reality Software and Technology(VRST), Mark Green (ed.), ISBN: 0-89791-825-8, p 163-172.

Slater, M., Steed, A. and Usoh, M. (1994) Steps and Ladders in Virtual Reality, ACM Virtual Reality Science and Technology (VRST), eds G. Singh and D. Thalmann, World Scientific, 45-54.

Slater, M., Steed, A., McCarthy, J., Maringelli, F. (1997) The Influence of Body Movement on Subjective Presence in Virtual Environments, Human Factors, in press.

Slater, M., Usoh, M. (1994) Body Centred Interaction in Immersive Virtual Environments, in N. Magnenat Thalmann and D. Thalmann (eds.) Artificial Life and Virtual Reality, John Wiley and Sons, 125-148.

Slater, M., Usoh, M. (1994) Body Centred Interaction in Immersive Virtual Environments, in N. Magnenat Thalmann and D. Thalmann (eds.) Artificial Life and Virtual Reality, John Wiley and Sons, 125-148.

Slater, M., Usoh, M., Steed, A. (1995) Taking Steps: The Influence of a Walking Metaphor on Presence in Virtual Reality, ACM Transactions on Computer-Human Interaction (TOCHI) 2(3) September, 201-219.

Slater, M., 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) 603-616, MIT Press.

Slater, M.,Usoh, M., Chrysanthou, Y. (1995) The Influence of Dynamic Shadows on Presence in Immersive Virtual Environments, Second Eurgraphics Workshop on Virtual Reality, M. Goebel, ed., Monte Carlo, Jan., 1995.

Smith, P. and Whetton, C. (1988) General Ability Tests (User's Guide) The National Foundation for Educational Research, ASE.

Song, D. and M. Norman (1993) Nonlinear Interactive Motion Control Techniques for Virtual Space Navigation, IEEE Virtual Reality Annual International Symposium (VRAIS), September 18-22, Seattle, Washington, 111-117.

Song, D., M. Norman (1993) Nonlinear Interactive Motion Control Techniques for Virtual Space Navigation, IEEE Virtual Reality Annual International Symposium (VRAIS), September 18-22, Seattle, Washington, 111-117.

TABLE 2.

Srinivasan, M.A., 1995, "Haptic Interfaces", In *Virtual Reality: Scientific and Technical Challenges*, Eds: N. I. Durlach and A. S. Mavor, pp. 161-187, National Academy Press.

Srinivasan, M.A., Basdogan, C., 1997, "Haptics in Virtual Environments: Taxonomy, Research Status, and Challenges", *Computers and Graphics*, (Special issue on "Haptic Displays in Virtual Environments"), Vol. 21, No. 4, pp. 393-404.

Steuer, J. (1992) Defining Virtual Reality: Dimensions Determining Telepresence, Journal of Communication 42(4), 73-93.

Strickland, D. (1996) A Virtual Reality Application with Autistic Children, Presence: Teleoperators and Virtual Environments, MIT Press, 5(3), 319-329.

Sutherland, I.E. (1965) The Ultimate Display, Proceedings of the IFIPS Conference, 2, 506-508.

Sutherland, I.E. (1968) Head-Mounted Three-Dimensional Display, Proceedings of the Fall Joint Computer Conference, 33, 757-764.

Synnott, A. (1993) The Body Social: Symbolism, Self and Society, Routledge: London and New York.

Teitel, M.A. (1990) The Eyephone, a Head Mounted Stereo Display, SPIE, 1256 Stereoscopic Displays and Applications.

Thorisson, K., Cassell, J. (1996) Why Put An Agent In A Human Body: The Importance Of Communicative Feedback in Human-Humanoid Dialogue, Abstract, in Lifelike Computer Characters '96. Snowbird, Utah, October 8-11, 1996, pp. 44-45.

Uno, S. and Slater, M.(1997) The Sensitivity of Presence to Collision Response, IEEE VRAIS 97, Sharon Stansfield, Larry Hodges and Mark Green eds, in press.

Vaananen, K. and K. Bohm (1993) Gesture Driven Interaction as a Human Factor in Virtual Environments - an approach with Neural Networks, in R.A. Earnshaw and M.A. Gigante (eds) Virtual Reality Systems, Academic Press, 93-106.

Vilhjálmsson, H.H. (1997) Autonomous Communicative Behaviors in Avatars, Masters Thesis, Learning and Common Sense Section, Media Laboratory, MIT <http://lcs.www.media.mit.edu/people/hannes/msthesis/>.

Ware, C. and S. Osborne (1990) Exploration and Virtual Camera Control in Virtual Three Dimensional Environments, Proceedings of the 1990 Symposium on Interactive 3D Graphics, ACM SIGGRAPH, March 1990, 175-183.

Welsh, R.B., Blackman, T.T., Liu, A., Mellers, B.A., Stark, L.W. (1996) The Effects of Pictorial Realism, Delay of Visual Feedback, and Observer Interactivity on the Subjective Sense of Presence, Presence: Teleoperators and Virtual Environments, MIT Press, 5(3), 263-273.

Wilson, P.N., N. Foreman (1993) Transfer of Information from Virtual to Real Space: Implications for People with Physical Disability, in M. Goebel (ed.) First Eurographics Workshop on Virtual Reality, 21-25.

Witmer, B.G., Singer, M.J. (1998) Measuring Presence in Virtual Environments: a Presence Questionnaire, Presence: Teleoperators and Virtual Environments, 7(3), June 1998, in press.

TABLE 2.

Zeltzer, D. (1992) Autonomy, Interaction and Presence, Telepresence, Presence: Teleoperators and Virtual Environments, 1, winter 1992, MIT Press, 127-132.

Zizek, S. (1991) An Introduction to Jacques Lacan Through Popular Culture, An October Book, The MIT Press, Cambridge Massachusetts, and London England.