A CAPACITY LIMITED, COGNITIVE CONSTRUCTIONIST MODEL OF VIRTUAL PRESENCE

David Nunez, MPhil.

Dissertation Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Computer Science

UNIVERSITY OF CAPE TOWN

December, 2007

Abstract

The Capacity Limited, Cognitive Constructionist (CLCC) model of presence is proposed as an information processing model of presence, which is demonstrated to have more theoretical power than extant models. The CLCC model defines information processing paths between attention, working memory, declarative memory and procedural memory, which operate to create and maintain a semantic context or bias. Bottom-up information entering the sensory cortices is filtered by attention into working memory where it forms temporary structures encoding the subject's experience of the VE. These structures also receive top-down information, which arises in declarative memory. This interaction of top-down and bottom-up data gives the entire model a semantic bias which attempts to keep the subject's construction of the environment semantically coherent. This allows for inferences and decision making, which translates into high presence. A semantically incoherent construction, or one which does not have enough working memory capacity allocated to it will result in poorer inferences about the environment, and reduced presence. If, as the CLCC model contends, presence involves information processing rather than simple perception, then one would expect to see working memory interference effects and semantic content effects in presence. Six studies were conducted to test these conjectures and validate the CLCC model. Studies 1 – 3 examined the role of working memory and attention on presence (the bottom half of the model), while Studies 4 – 6 examined semantic content and processing effects on presence (the top half of the model).

Study 1 manipulated working memory (WM) load during VE exploration. The CLCC prediction was that WM load would interfere with presence. Data from 177 subjects showed smaller effects than predicted: No WM effects on spatial presence, lower naturalness under spatial WM load, and lower engagement under verbal WM load. This suggests that spatial presence makes no use of WM, and that engagement and naturalness make limited use of it. While engagement seems to make use of semantic processing as predicted, naturalness seems to use spatial processing. Study 2 examined WM use by media decoders by repeating Study 1 with a text-based VE. Data from 114 subjects shows no WM effects exist on any of the four ITC-SOPI factors. This supports Study 1's finding that spatial presence does not use WM, but

contradicts results engagement and naturalness. Study 3 examined the relative contribution of attention and WM. 46 subjects viewed VE walkthroughs in three conditions: Viewing one walkthrough only (baseline), viewing two walkthroughs simultaneously (WM load condition), or viewing one walkthrough and a jumbled video simultaneously (attention load condition). The CLCC model predicted the WM load condition would interfere with presence the most, followed by the attention load condition, followed by the baseline. No difference was found across conditions, although naturalness and engagement predicted task performance, indicating that spatial presence is distinct from these factors, in agreement with the findings of Study 1 and 2.

Study 4 was a survey of semantic and processing effects on presence. Data from 101 computer gamers indicate that it is how often gamers play presence games (and not how many years they have been playing) that predicts how important they consider presence to their gaming experience. This suggests a moderate term activation effect rather than a long term learning effect. Furthermore, gamers with a high thematic inertia rate presence as important to gaming, indicating a processing effect. Finally, gamers who are capable of integrating non-diegetic music into their experiences rate presence as more important, which supports the CLCC notion that information processing of both semantic and perceptual information is important to presence. Study 5 followed up Study 4 by focusing on one specific content area. 461 flight simulation gamers completed the survey. Findings largely agree with those of Study 4, and strongly support the CLCC model prediction that highly specific expectations of content will reduce presence, while generalized expectations will increase it. Thematic inertia and priming were are also positively associated with presence, as predicted by the CLCC model. Study 6 manipulated non-diegetic information (background music) and semantic priming to test semantic processing in presence. The CLCC model predicted that all VE related information (semantic or perceptual) contributes to presence, particularly engagement and naturalness. 181 subjects were primed with materials semantically relevant or irrelevant to VE content, and then experienced the VE with no background music (baseline), music which semantically fit the VE, or VE music which was not a semantic fit. Priming did not influence presence as predicted, but non-diegetic music which fit the VE increased naturalness as predicted. The no-fit music produced the same presence scores as the baseline condition, indicating that it was filtered out by attention, as predicted by the CLCC model.

Overall, the CLCC model and data show that content effects occur in presence, and how these are mediated by declarative memory. It also shows that presence is a complex multi-level processing phenomenon. Spatial presence is at a cognitively low level, relying on perceptual (bottom-up) information, while engagement and naturalness are heavily dependent on conceptual (top-down) information, operating as a set of expectation-content comparisons which, when met by the content, lead to enhanced presence. These high and low cognitive forms of presence are largely independent, but do share some semantic effects, likely due to a reliance on common underlying cognitive processes such as priming and thematic inertia. The top half of the CLCC model (which encodes semantic meaning and explains content effects) is better supported that the bottom half (which predicted interference and attention effects). This finding is highly unexpected, as the literature on almost all extant models predicts an important role for attention in presence. From the data however, one must conclude that spatial presence makes no use of working memory, while cognitive higher forms of presence make use of limited amounts of working memory.

Acknowledgements

This was a large project which was not cheap to run. It would not have been possible without the financial support of these entities:

- The Collaborative African Virtual Environment System (CAVES), which was funded by the South African government's Department of Arts, Culture, Science and Technology.
- The South African National Research Foundation
- The Faculty of Humanities at the University of Cape Town

Many people supported an encouraged me during the project. Most prominent among them is Edwin Blake, with whom I have been researching presence for almost ten years now; he has as always had great faith in my insights, and has always been an enthusiastic supporter of my work. Sally Swartz, Colin Tredoux and Paula Ensor, who took an interest in my budding research career and offered me a research post and a lab to work in. Cara Winterbottom, Katherine Rix, Steven Bowles and Chomsky (my little web server) who assisted me in collecting data and running the seemingly endless line of subjects through the lab. The VEG research group, who patiently listened to and vetted many of the ideas in this dissertation. My labmates at the CVC lab with whom I shared sleepless nights during the dark days of CAVEAT development and testing, as well as many happy movie nights. Finally, Ilda Ladeira, who has to listen to my gripes (and the odd success) day in and day out for little more than scraps of cheese.

This project taught me a lot about research which you don't find in methodology texts. I subsequently found most of what I needed to know in this proverb by Antonio Machado:

Cuatro cosas tiene el hombre	*Four things man has*
que no sirven en la mar:	*which are of no use upon the sea:*
ancla, gobernalle y remos,	*anchor, tiller and oars,*
y miedo de naufragar.	*and fear of wrecking his ship.*

Table of Contents

ABSTRACT **3**

ACKNOWLEDGEMENTS **6**

CHAPTER 1: INTRODUCTION **11**

1.1 AIM OF THIS DISSERTATION 13
1.2 ASSUMPTIONS AND METHODS 13
1.3 OUTLINE OF THE DISSERTATION 14

CHAPTER 2: CONCEPTS AND MEASURES OF PRESENCE **16**

2.1 THE PHENOMENON OF PRESENCE 16
2.2 UNIFYING PRESENCE 18
2.3 IS PRESENCE A CONTINUOUS OR BINARY EXPERIENCE? 24
2.4 MEASURES OF PRESENCE 26
2.5 THE RELATIONSHIP BETWEEN CONCEPTS AND MEASURES 49
2.6 THE CONCEPT AND MEASURE OF PRESENCE USED IN THIS WORK 50

CHAPTER 3: A CRITICAL REVIEW OF CURRENT SIGNIFICANT MODELS OF PRESENCE 53

3.1 FIVE PROBLEMS FOR PRESENCE THEORY 54
3.2 THE FIVE PROBLEMS AS A YARDSTICK OF MODEL POWER 56
3.3 CURRENT SIGNIFICANT MODELS IN THE PRESENCE LITERATURE 57
3.4 CONCLUSION 108

CHAPTER 4: THE CAPACITY LIMITED, COGNITIVE CONSTRUCTIONIST MODEL OF PRESENCE (CLCC) **109**

4.1 STRUCTURAL BASIS 110
4.2 COGNITIVE CONSTRUCTIONISM 111
4.3 STRUCTURE AND COMPONENTS OF THE MODEL 113
4.4 PRESENCE IN THE MODEL 127
4.5 THE IMMERSION-PRESENCE RELATIONSHIP 131
4.6 FORMS OF PRESENCE BEYOND SPATIAL 132
4.7 HOW THE MODEL DEALS WITH THE FIVE PROBLEMS 133
4.8 CONCLUSION 136

CHAPTER 5: STUDY 1 - EFFECTS OF WORKING MEMORY LOAD ON PRESENCE **139**

5.1 THE WORKING MEMORY DUAL TASK PARADIGM 139
5.2 PREDICTIONS ABOUT WORKING MEMORY MADE BY THE CLCC MODEL 140
5.3 SAMPLE 142
5.4 APPARATUS 142
5.5 PROCEDURE 143
5.6 EXPERIMENTAL TASK 145
5.7 MEASURES 148
5.8 ANALYSIS & RESULTS 151
5.9 DISCUSSION 154
5.10 CONCLUSION 157

CHAPTER 6: STUDY 2 – WORKING MEMORY USE IN A LEARNED MEDIA DECODER 159

6.1 PREDICTIONS ABOUT WORKING MEMORY IN LEARNED MEDIA DECODERS 160
6.2 SAMPLE 161
6.3 APPARATUS 162
6.4 PROCEDURE 163
6.5 EXPERIMENTAL TASK 165
6.6 MEASURES 168
6.7 ANALYSIS & RESULTS 168
6.8 DISCUSSION 172
6.9 CONCLUSION 175

CHAPTER 7: STUDY 3 – RELATIVE CONTRIBUTION OF WORKING MEMORY AND ATTENTION TO PRESENCE....177

7.1 A DUAL-TASK METHOD FOR COMPARISON....177
7.2 PREDICTIONS ABOUT WORKING MEMORY, ATTENTION AND PRESENCE MADE BY THE CLCC MODEL 179
7.3 SAMPLE....180
7.4 APPARATUS....180
7.5 PROCEDURE....182
7.6 EXPERIMENTAL TASK....182
7.7 MEASURES....185
7.8 ANALYSIS & RESULTS....185
7.9 DISCUSSION....188
7.10 CONCLUSION....190

CHAPTER 8: STUDY 4 – AN EXPLORATION OF EXPERIENCE EFFECTS AND COGNITIVE PROCESSING IN PRESENCE....192

8.1 STUDYING CONTENT EFFECTS ON PRESENCE....192
8.2 CLCC MODEL PREDICTIONS....194
8.3 SAMPLE....195
8.4 PROCEDURE....195
8.5 MEASURES....196
8.6 ANALYSIS & RESULTS....198
8.7 DISCUSSION....202
8.8 CONCLUSION....204

CHAPTER 9: STUDY 5 – THE ROLE OF CONTENT KNOWLEDGE IN PRESENCE....206

9.2 PREDICTIONS OF CONTENT EFFECTS MADE BY THE CLCC MODEL....208
9.3 PROCEDURE....209
9.4 SAMPLE....210
9.5 MEASURES....210
9.6 ANALYSIS STRATEGY....213
9.7 RESULTS....214
9.8 DISCUSSION....218
9.9 CONCLUSION....225

CHAPTER 10: STUDY 6 – EFFECTS OF SEMANTIC CONTENT MANIPULATION ON PRESENCE 228

10.1 SEMANTIC PRIMING....228
10.2 NON-DIEGETIC MUSIC....229
10.3 PREDICTIONS ABOUT SEMANTIC CONTENT MADE BY THE CLCC MODEL....229
10.4 SAMPLE....230
10.5 APPARATUS....231
10.6 PROCEDURE....232
10.7 EXPERIMENTAL TASK....233
10.8 CONTENT MANIPULATIONS....233
10.9 MEASURES....236
10.10 ANALYSIS & RESULTS....236
10.11 DISCUSSION....242
10.12 CONCLUSION....246

CHAPTER 11: CONCLUSIONS....247

11.1 WORKING MEMORY AND ATTENTION IN PRESENCE....248
11.2 COGNITIVELY HIGHER-LEVEL AND LOWER-LEVEL FORMS OF PRESENCE....249
11.3 ‘PRESENCE’ OR SPATIAL PRESENCE?....252
11.4 THE ROLE OF EXPECTATION IN PRESENCE....253
11.5 FURTHER DEVELOPMENT OF THE CLCC MODEL....254

APPENDIX A: THE INDEPENDENT TELEVISION COMMISION'S SENSE OF PRESENCE INVENTORY (ITC-SOPI)....256

A.1 OVERVIEW 256
A.2 PRESENTATION OF ITEMS....257
A.3 ITEMS IN EACH FACTOR 257

APPENDIX B: THE DIFFERENTIAL EMOTION SCALE, SECOND EDITION (DES-II) 260

B.1 260
B.2 PRESENTATION 260
B.3 ITEMS 260

APPENDIX C: GAMER'S SURVEY USED IN STUDY 4....261

C.1 PRESENTATION OF ITEMS....261
C.2 ITEMS IN EACH FACTOR 261

APPENDIX D: FLIGHT SIMULATOR USERS' SURVEY USED IN STUDY 5 265

D.1 PRESENTATION OF ITEMS....265
D.2 ITEMS IN EACH FACTOR 265

APPENDIX E: VE SCENARIO INSTRUCTIONS 270

E.1 PRESENTATION OF INSTRUCTIONS....270
E.2 HOSPITAL SCENARIO....270
E.3 MONASTERY SCENARIO 270

APPENDIX F: PRIMING MATERIALS USED IN STUDY 6 271

F.1 OVERVIEW....271
F.2 PRESENTATION 271
F.3 MONASTERY-RELEVANT SLIDES (VE-RELEVANT)....271
F.4 MONASTERY-IRRELEVANT SLIDES (VE-IRRELEVANT) 275

APPENDIX G: SUMMARY TABLE OF MODELS REVIEWED IN CHAPTER 3....277

APPENDIX H: SUMMARY TABLE OF CLCC MODEL EVIDENCE....277

REFERENCES....285

Table of Figures

FIGURE 3.1: THE TWO-POLE MODEL. 58
FIGURE 3.2: THE ENVIRONMENT SELECTION MODEL.. 60
FIGURE 3.3: THE THREE-POLE MODEL. 72
FIGURE 3.4: THE FOCUS LOCUS SENSUS (FLS) MODEL. 80
FIGURE 3.5: THE LEVELS OF PRESENCE (LOP) MODEL. 82
FIGURE 3.6: THE FIRST PHASE OF THE MEC MODEL (SSM CONSTRUCTION). 95
FIGURE 3.7: THE SECOND PHASE OF THE MEC MODEL. 98
FIGURE 4.1: THE STRUCTURE OF THE CLCC MODEL. 115
FIGURE 5.1: IMAGE FROM THE TRAINING VE 144
FIGURE 5.2: IMAGE FROM THE HOSPITAL ENVIRONMENT 144
FIGURE 5.3: IMAGE FROM THE HOSPITAL ENVIRONMENT 145
FIGURE 5.4: PANEL PROVIDING THE VERBAL CODE 147
FIGURE 5.5: LOCK FOR THE VERBAL CODE 148
FIGURE 5.6: THE PANEL SHOWING THE SPATIAL CODE. 150
FIGURE 5.7: THE LOCK FOR THE SPATIAL LOAD CONDITION. 150
FIGURE 5.8: TASK PERFORMANCE AS A FUNCTION OF WORKING MEMORY LOAD, SHOWING THE EFFECT OF THE EXPERIMENTAL MANIPULATION 152
FIGURE 6.1: THE TEXT-BASED TRAINING ENVIRONMENT. 164
FIGURE 6.2: THE MAIN HOSPITAL VE IN A ROOM WITH NO LOCK AND NO PANEL. 165
FIGURE 6.3: A PANEL SHOWING A CODE OF LENGTH 5 FOR THE VERBAL LOAD CONDITION 167
FIGURE 6.4: THE LOCK IN THE VERBAL LOAD CONDITIONS 167
FIGURE 6.5: SCATTERPLOT OF EFFECTIVE WORKING MEMORY LOAD VERSUS TASK PERFORMANCE 170
FIGURE 7.1: SAMPLE OF ONE OF THE INSTRUCTION SLIDES 181
FIGURE 7.2: SAMPLE OF INSTRUCTIONS (FOR THE ATTENTION CONDITION) 181
FIGURE 7.3: SCREENSHOT OF THE BASELINE CONDITION (ONE VIDEO ONLY). 183
FIGURE 7.4: SCREENSHOT OF THE WORKING MEMORY CONDITION. 183
FIGURE 7.5: SCREENSHOT OF THE ATTENTION CONDITION 184
FIGURE 7.6: AVERAGE NUMBER OF LEFT KEY PRESSES BY CONDITION 186
FIGURE 9.1: CONTENT FACTORS, IN DECREASING LEVEL OF GENERALITY OF EXPECTATION PLOTTED AGAINST THEIR PARTIAL CORRELATION WITH THREE ITC-SOPI FACTORS 226
FIGURE 10.1: IMAGE FROM THE MONASTERY ENVIRONMENT (UPSTAIRS LANDING) 232
FIGURE 10.2: IMAGE FROM THE MONASTERY ENVIRONMENT (COURTYARD) 233
FIGURE 10.3: A BOOK CENTERED ON THE SCREEN ABOUT TO BE COLLECTED 234
FIGURE 10.4: SAMPLE OF THE VE-IRRELEVANT PRIMING MATERIAL 235
FIGURE 10.5: SAMPLE OF THE VE-RELEVANT PRIMING MATERIAL 235
FIGURE 10.6: HISTOGRAM OF TASK PERFORMANCE (NUMBER OF OBJECTS COLLECTED) 237
FIGURE 10.7: MEANS PLOT FOR THE NON-DIEGETIC MUSIC CONDITION EFFECT ON NATURALNESS. 242

Chapter 1: Introduction

Presence, the sense that one is experiencing a mediated or virtual environment as if it were real and non-mediated, is strongly associated with mediation and technology, but in the end there is always a person, with intelligence, memories, biases and expectations, who takes in the stimuli produced by the environment, and has (or fails to have) a presence experience. This dissertation is about the role of the subject in presence. People are not passive agents in the world – they process what they experience, and they act based on what they know. This dissertation aims to uncover how much of the presence experience is attributable to this processing and knowledge. It is probably best to begin by giving a little history to put this research in context. Given the topic of this dissertation (virtual presence), one might expect this introduction to say something about Ivan Sutherland's "perfect display" from the late 1960s (Sutherland, 1965), or Marvin Minsky's early work on on telepresence in 1970s (Minsky, 1980); the kind of thing one might call a "standard history" of virtual reality and presence. Such histories are interesting, but they do not reflect what this dissertation is about at all, because being histories of science and engineering, they place too much emphasis on presence as an aspect of a technology. Rather, I would like to set a broader context, free of reference to any particular technology, by beginning this historical excursion a little earlier than the 1960s. I would like to begin in the mid 1660s.

During this time, Japanese poets were discovering a new medium, the *haiku* (Ueda, 1996). Much like virtual environments (VEs) which evolved from television and film, haiku evolved from an earlier, well respected and well established poetic form called *tanka* (Ueda, 1996). A *tanka* aims to evoke a moment or emotion chosen by the poet, within the constraints of only thirty-one syllables (Ueda, 1996). *Haiku*, however, was far more ambitious. Being largely defined by minimalist-minded Zen Buddhist monks such as Matsuo Basho (1644 – 1694), all *haiku* aimed to evoke one particular emotion – *sabi* – using only seventeen syllables. Basho himself became quite proficient at evoking *sabi* within the difficult technical constraints of *haiku* – consider these two examples of his work:

Moonlight slanting	*Another year is gone;*
through the bamboo grove;	*and I still wear*
a cuckoo crying.	*straw hat and straw sandal.*

(Ueda, 1982)

What exactly is *sabi*? *Sabi* was difficult to define, even by the *haiku* masters of that era. In modern Japanese, the word has transformed into *sabishii*, and refers to a bittersweet, peaceful loniless (Ikeno & Davies, 2002). For the *haiku* masters and their disciples, *sabi* was a complex emotion, and much discussion was devoted to defining it and understanding its nature. One of the most famous definitions of *sabi* was given by one of Basho's disciples, Kyoraisho. During a discussion group, Kyoraisho is quoted as having said

> *sabi is the colour of haiku;*
> *it is different from tranquility.*
> *For example, if an old man dresses up in armour and helmet*
> *and goes to the battlefield,*
> *or in colorful brocade kimono, attending his lord*
> *at a banquet, sabi is like this old figure.* (Ueda, 1982)

Perhaps *sabi* is a concept not too far from presence. Much like *sabi*, we who study presence struggle to define and capture its nature, although anyone can identify it when they experience it. Like *sabi* in *haiku*, presence is the colour of virtual reality (if one can paraphrase Steuer's 1992 definition of presence that way). Like *sabi* which can be evoked using only seventeen syllables, very little external stimulation is required to bring about presence (as little as a few polygons on a screen; Slater *et al.*, 1995c), but at the same time it is difficult to arrange stimuli so as to guarantee a good presence experience. The most fundamental and important similarity between *sabi* and presence, however, is that they are both emergent properties of an information processing system which includes a human subject and a mediated space. The stimuli in *haiku* (the painstakingly selected seventeen syllables), much like those of a VE, need to be perceived by a subject, and decoded; then, as these stimuli are interactively processed in the context of the subject's previous experiences and knowledge of the world, the desired experience (be it *sabi* or presence) emerges. This dissertation is my attempt to convince the reader that this simple explanation of presence is true.

1.1 Aim of this dissertation

The past five years have seen significant advances in presence theory. Several important predictive models have appeared (see chapter 3), and there has been significant improvements in thinking about the presence phenomenon itself. Theorists such as Biocca (2003), Timmins and Lombard (2005), Lee (2004) and Slater (2002) have identified phenomena and constraints to presence theory which allow for more exact development and evaluation of models (see chapters 2 and 3). It is from these developments that this dissertation begins. It aims to develop a model of presence which is able to satisfy all of the following criteria:

1. Can explain presence in immersive as well as non-immersive media (Biocca, 2003; Lombard & Ditton, 1997)
2. Is equally suitable for explaining behaviour in real and virtual environments (Biocca, 2003; Timmins & Lombard, 2005)
3. Is consistent with current models of cognition, and can be explained as an evolved natural process (Biocca, 2003; K. M. Lee, 2004)
4. Can explain perceptual as well as content effects (Lessiter *et al.*, 2001; Robillard, 2003)
5. Can explain the duality of presence as a binary (Slater, 2002) and continuous phenomenon (Wirth *et al.*, 2007)
6. Is well supported empirically

Although a number of useful models of presence already exist (see chapter 3), none are able to meet all the constraints listed above. By careful examination of each extant model, this dissertation will define the Capacity Limited, Cognitive Constructionist model of presence (CLCC), conceptually show that it has more predictive power than extant models, and empirically validate its central aspects using six studies.

1.2 Assumptions and methods

This dissertation begins with the assumption that presence is a psychological phenomenon which occurs due to the interaction of a subject with external stimuli, which may or may not be related to a mediated environment (after Biocca, 2003; K. M. Lee, 2004). Following this assumption, the methods used are those of computational cognitive psychology; the subject's mind is assumed to be an information processing system which can be modeled to produce specific predictions which can be verified empirically (after Neisser, 1967). By and large, these assumptions and methods are non-controversial in the presence field, although this dissertation places far more emphasis on the role of information content than previous work (this is justified in chapter 4). The measurement of presence in all studies is done by the use of self-reports, which have been at the center of some debate (see 2.4.6 in chapter 2). This choice was made after a review of the psychometric properties of various self-report and other measurement methods, and is justified in section 2.6 (chapter 2).

1.3 Outline of the dissertation

This dissertation is divided into two parts. Part I (chapters 2 – 4) presents the capacity limited, constructionist cognitive (CLCC) model of presence, with its theoretical background. Part II (chapters 5 – 10) presents six studies which were used to empirically evaluate the CLCC model:

1.3.1 Outline of Part I

Chapter 2 presents a review of major conceptualizations of presence, and some efforts to unify theoretical strands. It then moves on to a review of measurement techniques, and concludes by selecting a definition and measure of presence to be used for the remainder of the dissertation. Chapter 3 reviews major presence models, and evaluates them in terms of theoretical soundness, empirical evidence, and their ability to deal with major theoretical problems in presence. Chapter 4 presents the CLCC model, and evaluates it theoretically, using the same criteria applied in chapter 3.

1.3.2 Outline of Part II

Chapters 6 and 7 present two working memory loading studies which examine presence under a dual-task paradigm. Chapter 7 uses a similar paradigm, but examines presence under divided attention. Chapters 8 and 9 present two survey conducted on computer gamers to examine the role of experience, content expertise and other processing factors in presence. Chapter 10 follows this with an experiment which examines the role of semantic coherence of the VE on presence. Chapter 11 presents conclusions for the dissertation as a whole.

Part I:

The CLCC model and its theoretical foundations

This part of the dissertation presents the capacity limited, cognitive constructionist model of presence as a theoretical entity. Chapter 2 begins by examining the concept of presence as it currently exists in the literature, and derives a flexible and powerful concept of presence to model. Measures of presence are then considered in order to find a valid and reliable basis for the empirical evaluation of the CLCC model in Part II. Chapter 2 critically reviews the most influential current models of presence in terms of their explanatory power and degree of supporting empirical evidence. Finally, Chapter 3 presents and justifies the CLCC model, and compares it, on a theoretical basis, with the extant models of presence reviewed in Chapter 2.

Chapter 2: Concepts and measures of presence

Currently, two significant theoretical hurdles stand in the way of presence research. First, although the presence experience is easily identifiable for the subject, a common scientific definition remains elusive; and second, little agreement exists in the field about how presence should be measured. Any empirical work must find some way to overcome these problems in order to claim some form of theoretical validity. This chapter has two goals: First, to critically review the current concepts of presence while focusing on efforts to unify them into a global concept, and argue for the importance of a cognitive perspective in unification; and second, to evaluate current presence measures so as to select a measure for the empirical component of this project which is consonant with the chosen presence concept.

2.1 The phenomenon of presence

Presence is a phenomenon which is associated with being a user of a system which mediates an environment. One definition which is a useful starting point for discussion is that provided by the International Society for Presence Research (ISPR), which is an international association of mostly academic presence researchers. That definition, which is heavily influenced by the synthesis of Lombard and Ditton (1998) is:

> *Presence (a shortened version of the term "telepresence") is a psychological state or subjective perception in which even though part or all of an individual's current experience is generated by and/or filtered through human-made technology, part or all of the individual's perception fails to accurately acknowledge the role of the technology in the experience. Except in the most extreme cases, the individual can indicate correctly that s/he is using the technology, but at some level and to some degree, her/his perceptions overlook that knowledge and objects, events, entities, and environments are perceived as if the technology was not involved in the experience. Experience is defined as a person's observation of and/or interaction with objects, entities, and/or events in her/his environment; perception, the result of perceiving, is defined*

as a meaningful interpretation of experience. (International Society for Presence Research, 2000, paragraph 1)

This standardized sounding definition is actually a little misleading. In practice, pinning down exactly what presence is has been made far more difficult due to a proliferation of subtypes of presence. Here are some examples:

- Social presence: When subjects interact with simulated characters and mediated users as if they were not mediated (K. M. Lee, 2004)

- Co-presence: When the subject feels as if other mediated users are really in the same space as themselves (Slater *et al.*, 1999)

- Story presence: When subjects feel as if they are inside the events of a story (Brown *et al.*, 2003)

- Cognitive presence: When subjects take the VE as the basis for cognition rather than the non-mediated world (Nunez & Blake, 2001)

- Relational presence: When two subjects who experience a mediated relationship feel an emotional connection (Maguire & Connaughton, 2006).

- Spatial presence: Where subjects feel as if they are occupying the virtual space, and they are sharing the space with the objects in it (S. Lee *et al.*, 2004b) – more confusingly, this is often simply referred to as 'presence' (Slater, 2003a), which gives the impression that it is somehow the top of a hierarchy of presence concepts.

This is by no means an exhaustive list - Lombard and Jones (2006) are in the process of compiling a definitive taxonomy. A more constructive way to understand presence might be to consider all definitions as instances of a broader principle. In essence, presence arises when one person interacts with an environment (which may or may not be populated). This interaction could be mediated, or unmediated (Biocca, 2003;

IJsselsteijn, 2002; K. M. Lee, 2004; Lombard & Ditton, 1997). This very basic notion of presence originates from reports of operators of remote systems who, while working with their mediated systems, sometimes felt as if they were located at the remote site rather than at the workstation (Minsky, 1980). This phenomenon was described in detail by Sheridan (1992a), who suggested that presence could be applied to a user of a virtual environment, as this was simply a special case of teleoperation where the remote site was virtual rather than real. This extremely broad definition ("a feeling of being at the remote or virtual site rather than at the workstation") has continued to be extremely influential. This basic concept has never truly been replaced, but rather refined. (a few notable exceptions exist – see Floridi, 2005 for an example of a different but not widely adopted notion of presence based not on the sense of being at a remote site, but on the notion of a mismatch between perceptual sensors at the local and remote sites).

2.2 Unifying presence

Sheridan's simple concept may have been influential, but ultimately it was too nebulous to base a program of empirical research on. This led to a need for something more specific, and is probably what sparked the creation of new 'types' of presence. The creation of new definitions of presence solved the specificity problem, but created another: Along with new definitions came new measures, and comparing findings between studies became progressively more difficult (Kalawsky, 2000). Attempts were made to unify these concepts under a smaller number of umbrella concepts. For instance, Schloerb (1995) created the categories *subjective presence* and *objective presence* to simplify the picture (Schuemie *et al.*, 2001), but these were effectively categorizations of presence *measures* rather than presence *concepts*, and did not alleviate the problem. A similar but more substantial categorization was proposed by Heeter (1992), which operates on a level of analysis principle: Presence can be *personal* (the subject feels that they are in the virtual environment), *social* (the subject feels that the other characters and avatars in the virtual environment are real social actors) or *environmental* (the environment responds and reacts to the subject). As with Schloerb's effort, this is more of a methodological advance than a theoretical one. The categories help with deciding which measures and events are of interest, but in an overly reductionistic way. Consider a simple on-line gaming session as a common example of a virtual environment experience. Here players may experience personal presence (they feel present in the space), and simultaneously, social presence (as they work in teams to achieve the game's goals), and environmental presence (as they see doors open and objects tumble as they bump into them). It is methodologically useful to say that this subject is experiencing environmental or social presence, as it clears up what the variables of interest are; but from a theoretical perspective, it does not provide much insight, particularly because it forces the assumption that these forms of presence might be phenomenologically different beyond simply being different levels at which the phenomenon can be analyzed.

An important move towards a theoretically useful and empirically workable definition was published by Lombard and Ditton (1998). They began by reviewing both the theoretical and empirical literature. This included sources not only from the field of VE research, but from communications and media research as well. From this review, they created six categorizations for both explicitly stated and implicitly held formulations of presence:

1. *Presence as social richness* – presence arises when subjects perceive the environment as warm, sensitive, personal, intimate and immediate (these definitions typically arise from communications researchers).

2. *Presence as realism* –presence occurs when the medium is able of reproducing realistic objects, characters and events (these definitions arise mostly from human factors and systems engineers)

3. *Presence as transportation* – subjects experience either a sense of being in another place, that objects have been transported to share the same space as the subject, or that a group of subjects have been transported to the same place (these definitions arise from various sources, including communications research and literary theorists).

4. *Presence as immersion* – the medium becomes the primary source of sensory data for subjects (perceptual immersion), which may leads to a sense of involvement in the medium (psychological immersion). These definitions are often used by virtual reality and immersive media researchers.

5. *Presence as social actor within the medium* –subjects interact with characters in the medium as if they were social actors (these definitions are found mostly in the communications and shared spaces fields).

6. *Presence as the medium becoming a social actor* – subjects interact with the medium itself as if it were a social actor (these definitions are used mostly in the fields of robotics and human-computer interaction).

This ambitious project aimed to unify not only the presence concepts which cover what Heeter would call personal presence, but also those which deal social presence – a feat which has not been replicated since, and no doubt has led to this being one of the most widely cited papers in the field. Lombard and Ditton went further than simply organizing the definitions into six categories. They examined the categories searching for a common, unifying factor. They concluded that the factor is *the perceptual illusion of non-mediation*. They argue that in all six categories, presence occurs because the subject is responding to environment or characters as if the system (or more specifically, the technological aspects of the system) were not there. This occurs either because the system becomes a sort of window into the environment (in the classic 'ultimate display' sense proposed by Sutherland, 1965), or because the system is perceived as having transformed itself into a social actor (Lombard & Ditton, 1997).

An interesting consequence of Lombard and Ditton's definition is that it shifts the focus of presence onto perception. Much of the early examples provided by, for example, Sheridan (1992a; , 1992b), Ellis (1996) and Slater (1993a; , 1995c), had an underlying assumption that given the correct conditions, presence would occur automatically (part of the *presence as immersion* family of definitions). However, by highlighting perception rather than sensation, Lombard and Ditton suggest that, like perception, presence is a constructive, active process where the subject combines aspects of their own experience with the stimuli arising from the medium to create the experience for themselves:

> *Because it is a perceptual illusion, presence is a property of a person. However it results from an interaction among formal and content characteristics of a medium and characteristics of the media user, and therefore it can and does vary across individuals and across time for the same individual.* (Lombard & Ditton, 1997)

The idea that cognition must play an important role in presence has slowly gained popularity. Heeter has argued that presence can only arise from an interaction of top-down and bottom-up data, and furthermore, can only occur in a meaningful environmental context (Heeter, 2003). Nunez and Blake (2001) re-examined Lombard and Ditton's unifying principle using a cognitive perspective, with the aim of combining explanations of presence as a sensation (such as used by Sheridan, 1992a) with behavioural presence (for example, the notion of behavioural realism proposed by Freeman *et al.*, 2000). This approach (which was termed 'cognitive presence', misleadingly suggesting a new type of presence rather than an omnibus concept), cognition is argued to be the appropriate level of analysis, as all environments must be decoded and processed before any presence or action can occur. Cognition is therefore seen as the bottom-most level at which presence can be analyzed (Nunez & Blake, 2001). It should be noted that like Lombard and Ditton, Nunez and Blake do not present any discussion of the relative difficulty of using this simple level of analysis to explain the more complex forms of presence (such as social presence), and its usefulness in these domains is therefore unknown.

Cognitive presence as an adaptation of Lombard and Ditton's principle has only seen a limited amount of acceptance in the literature (used by, for instance, Hwang *et al.*, 2004; S. Lee *et al.*, 2004a). Rather, it is the examination of presence as a neuropsychological phenomenon which is showing signs of being a major unifying principle for presence. IJsselsteijn (2002) argues that presence should be explained as a multi-level phenomenon, with the bottom-most level being neural processing; then cognitive, phenomenological, social and other explanations should be considered as progressively complex levels which remain consonant with lower levels (IJsselsteijn, 2002). This view is perhaps an improvement over the monolithic approaches to presence unification proposed by Lombard and Ditton and Nunez, and Blake, as they allow switching to a particular level of explanation as required by the research problem without violating the principles of the more fundamental levels of explanation. Unfortunately, the undoubtedly frightening complexities of how a neural theory of presence might mesh with a social theory are not discussed. IJsselsteijn completes his argument by speculating on future research directions for a neural understanding of presence, including possible physiological correlates of presence (some of which are later echoed by Sanchez-Vives & Slater, 2005). Some of these ideas have been examined empirically (see 3.3.3.3 in chapter 3).

The most recent move towards unifying presence (which again comes from a cognitive perspective) is by Lee (2004). It is the first paper to substantially address possible ultimate as opposed to proximate causes of presence (Biocca, 2003 had previously noted the importance of considering evolutionary processes in presence). Lee argues that the interesting feature of presence is not that it occurs, but it sometimes fails to occur, given that it is an automatic reaction to particular stimuli (K. M. Lee, 2004). He therefore opposes presence concepts which emphasize the role of subjects' own conscious efforts to bring about the experience (such as the "suspension of disbelief" argument used by Slater & Usoh, 1993b). The mental architecture which Lee uses to support his argument is a cognitive modular architecture (Fodor, 1983; Tooby & Cosmides, 1990) which has recently been popularized by Pinker (1997; , 2002). In this architecture, the mind consists of various modules which are specialized to take in very particular types of data, transform them, and output to other similar modules. Each module is fairly simple, and optimized by natural selection for its particular purpose (Tooby & Cosmides, 1990). Because these modules are simple, they do not have complex filters for incoming data – they will respond even to stimuli which only resemble the intended objects; they thus respond to mediated scenes in the same way as they do to real scenes (K. M. Lee, 2004). For instance, the face processing module will respond to facial expressions even if the faces are depicted as simple cartoons rather than real faces (Heraz & Frasson, 2006). Of course, the mediated stimuli must fall within certain limits for the modules to take over their processing, hence the general finding that scenes with more fidelity lead to more presence (K. M. Lee, 2004). Although the modules process automatically, they can still be controlled to some extent by the central executive (Pinker, 1997), This is not however done by the willful *activation* of modules (as the "willing suspension of disbelief argument would contend), but rather by the willing *inhibition* of modules so that other reactions and behaviours can be expressed instead (Fox *et al.*, 2005).

Lee (2004) proposes that two types of presence can arise from this mechanism:

1. *Spatial presence through the folk physics module:* The folk physics module, which processes motion, three-dimensional spatial data and basic physics (Pinker, 1997) reacts automatically to sensory stimuli to form cognitive maps, mental models and other structures which are then handed off to other processes for the selection of appropriate action (Tooby & Cosmides, 1990). As these modules respond to simple features of sensory inputs, they can be easily fooled by mediation technologies (K. M. Lee, 2004). For instance, the spatial processing component of vision makes use of a few simple features such as relative size and occlusion to extract a three-dimensional space from a two-dimensional retinal image (Tong, 2002). Lee (2004) argues that most of the published findings which report on immersion effects on presence occur because the technology is capitalizing on the automatic interpretation of these features by the folk physics module. When this occurs, the subject will respond to the displayed scene in the way they react to unmediated objects, and probably experience a sense of spatial presence in doing so.

2. *Social and co-presence through the folk psychology module:* The folk psychology module is used to infer mental states in others (Baron-Cohen, 1995). It takes as input semantic information from the speech processing module, visual information and facial expression data from the face recognition and processing modules (Pinker, 1997). As with the folk physics module, it is activated by a few key features, and can be fooled by a display (K. M. Lee, 2004). The module is used to initiate social behaviours and make psychological predictions about other people. When activated, it leads to a sense that the object being interacted with has intelligence or personality (Baron-Cohen, 1995). As with the folk physics module, an active folk psychology module leads to subjects responding as they do to non-mediated persons, and a sense of social or co-presence could result.

The impressive advantage that Lee's unifying structure has over that of Lombard and Ditton (1998) is that because modules are specialized for particular tasks, he can retain the "typed" flavour of the different presence definitions (allowing spatial and social presence, for instance, to retain their own phenomenological distinctness), while at the same time explaining them using a single principle (i.e. the automatic processing by evolved modules). Nonetheless, for all the differences in the unifying attempts described above, it is interesting that all make use of cognitive explanations (or more generally, human information processing explanations) to produce theoretically reasonable umbrella terms.

2.3 Is presence a continuous or binary experience?

Until the year 2000, presence was generally considered to be a psychological variable, similar to mental workload (Sheridan, 1992a) and situational awareness (Prothero *et al.*, 1995). This led to the assumption that presence existed on a continuous scale, from highly present in the environment of interest, to totally absent and unaware. Perhaps the most compelling recent treatment of presence as a continuous phenomenon comes from Heeter (2003), who argues not only that presence fluctuates from moment to moment within a given environment, but also that individuals differ in terms of their presence experiences of the real world. The notion of presence as continuous is so convincing, that it has been carried through into most models and theories of presence (Biocca, 2003; Riva & Waterworth, 2003; Slater *et al.*, 1994; Waterworth & Waterworth, 2001; Wirth *et al.*, 2007). As all current models of presence are written at a conceptual rather than strictly predictive level, the question of whether presence is continuous or binary is not particularly critical – most models could theoretically cope with both. It is in the area of measurement that the question becomes critical.

Early measures of presence included methods which gave both quantified estimates on a scale and dichotomous ratings. For example, Barfield and Weghorst (1993), Hendrix *et al* (1996a; , 1996b) as well as Witmer and Singer (1998) used questionnaires which asked subjects to rate their experience of 'being there' (to varying degrees of sophistication) on Likert or semantic differential scales. On the other hand, Schloerb (1995) suggested that subjects could be shown a view through a window or peep-hole, which could either be simulated or real; and the more often the subject misidentified the simulated scene as real, the more present the subject could be said to be. Similarly, Hoffman *et al.*(1995) suggested that subjects could be asked to recall if the event occurred in a virtual space or real space; errors in favour of selecting the real space were taken as an indication of presence. It should be noted that these techniques of binary measurement (present/not present) did not make assumptions about the nature of the construct they were measuring. They simply used a binary quantification of what could well be a continuous experience.

That the presence experience is a dichotomous phenomenon was indirectly argued by Lombard and Ditton (1998). For them, presence is a perceptual illusion, and so it must either occur or not (interestingly, since then Lombard and Ditton have published a validated continuous measure of presence, which implies they currently see it as a continuous phenomenon – see Lombard & Ditton, 2004). Currently, the most obvious manifestation of presence as a binary concept is that proposed by Slater (2002). In this view presence is the outcome of a choice made at a perceptual level by the subject (see 1.3.1.1 in chapter 3). Subjects are therefore in one environment or the other, and the change between them is not gradual, but a sudden 'break' experience, which subjects are capable of reporting (Slater & Steed, 2000). Slater argues that presence is a Gestalt, and therefore cannot be graded. By analogy, he compares presence to reversible ambiguous figure illusions (see for instance Girgus *et al.*, 1977), which are also considered to be Gestalts. In these illusions, a stimulus figure seems to switch in meaning (one moment appearing as a rabbit, the next as a duck, for instance), but is never perceived as anything in between. He further argues that the 'break in presence' phenomenon would not exist if the experience were graded. While convincing, Slater's argument is missing two parts: First, while there is no doubt that presence can end by a break (this is documented in Slater & Steed, 2000), it is not clear that breaks are the only mechanism by which presence ends. In fact, Slater and Steed argue that subjects become presence in a VE by a different mechanism which does not lead to a break sensation (Slater & Steed, 2000). But, in order to become present in the VE, the subject must first stop being present in the real environment, which implies that there are at least two ways one can stop being present (see 4.4.2 in chapter 4). It is possible that the presence experience is in fact continuous, but in a 'break in presence' situation, the speed at which the change occurs makes it seem as if it is an instantaneous break. Secondly, Slater does not provide a proper role for attention in his dichotomous model of presence (Slater, 2002), especially given that many breaks in presence occur because of the sudden appearance of distracters (such as rendering glitches, interference from cables, etc.). Models of presence which do take into account modern findings on attention (see for instance Lombard & Ditton, 1997; Waterworth & Waterworth, 2001; Wirth *et al.*, 2007), conclude that because attention can be divided, and does allow for a limited degree of simultaneous processing (Baddeley, 1986; Treisman, 1969), presence should be possible to varying degrees based on the amount of attention that is available for processing the scene. Of course, because large changes in the stimuli field can lead to the sudden selection of a particular subset of stimuli for processing (Treisman *et al.*, 1992), the break in presence phenomenon can also be explained from a divided attention perspective. It is therefore highly likely that presence, if it is indeed as strongly related to attention as the literature claims, is a continuous phenomenon. The fact that it could be continuous, of course, neither precludes nor invalidates *measures* of presence which quantify it into two levels.

A final possibility to explain the apparent contradiction between models and measures in this regard is that presence is in fact a binary phenomenon, and that during any one VE experience, a subject will experience a large number of drifts into and interruptions out of presence. If one is using a binary measure of presence (such as BIPS discussed in 2.4.5 below), then it will seem as if the experience is binary, as each interruption will be reported. However, if a continuous measure is used (such as the TPI discussed in 2.4.1.6 below), it is possible that the subject gives an 'averaged' impression of the experience, interpreting more interruptions to mean less presence when asked to place it on a Likert scale. This would then give the impression that presence is continuous. Given the current empirical data, it is not possible to falsify this notion, and it must remain as a possible (and likely) explanation for the apparent contradiction in the literature.

2.4 Measures of presence

The second great debate in presence is that around measurement. Generally speaking there are three forms of presence measurement – the first two, self-report scales (Lessiter *et al.*, 2001; Slater *et al.*, 1994) and behavioural observation (IJsselsteijn, 2004; IJsselsteijn *et al.*, 2000), have been discussed since the early 1990s, but the third, breaks in presence (Slater & Steed, 2000), is a new and interesting addition to the literature. Any measure can have two uses: one is simply to measure presence in an empirical setting (Nowak *et al.*, 2006; for instance as is done by Robillard, 2003; Vinayagamoorthy *et al.*, 2004). Another, less familiar use is to use the measure to generate a theory of the dimensional structure and organization of presence by the use of factor analysis (as has been done by Lessiter *et al.*, 2001; Witmer *et al.*, 2005). This particular use has been fairly controversial, leading Waller and Bachman (2006) to criticize the technique on the following grounds:

1. Factor analysis (and principal components analysis) are essentially descriptive. They simplify data by projecting it onto a space (Neter *et al.*, 1988). Effectively, a factor analysis re-organizes the data, with some associated loss of variance; it is not an inferential technique and is therefore not capable of testing hypotheses (Neter *et al.*, 1988). It is not appropriate to use these techniques to test theories of the structure of presence (with one exception – see point 3 below).

2. Many decisions required by factor analysis are made *a priori.* These include the items to include in the analysis, the type of rotation applied to the data before projection, and (most importantly) the number of factors to project to (Waller & Bachmann, 2006). Although one can measure how well the data fits the current factor structure (by considering the factor eigenvalues and loadings - Neter *et al.*, 1988), this would only reflect that the data fits the number of factors chosen. If the choice of optimal number of factors scores could be objectively made, then it might be an appropriate method for testing theories, but at best all it can do is provide an estimate of how well a particular number of factors fit a particular sampled data set.

3. Although factor analysis is limited, it has some use. Rather than using the exploratory factor analyses currently in the literature (Waller & Bachmann, 2006), one can employ *confirmatory factor analysis*, where one begins with a specified factor structure (generated from a theory), and the data set is fitted to this structure. One can thus use this technique to compare theories (Waller & Bachmann, 2006). This is done by fitting all competing theories to a given

data set, and seeing which fits with the smallest error (of course, this is assuming that the data has a large degree of external validity, and that the sample is large enough to rule out idiosyncratic artifacts within the data).

Factor analysis is a popular technique in presence research, and it has largely gone unopposed as a method up till the publication of Waller and Bachmann's paper. Although their arguments are correct, it seems that they may have taken the argument too far. First, the way in which factor analysis is being used by at least some presence researchers (such as Lessiter *et al.*, 2001; Wirth *et al.*, 2007) is indeed descriptive. The complexity of the data is being reduced into a manageable number of concepts. In most cases, the factor analyses show that the factors are highly inter-correlated, which is interpreted by some as evidence that there is either a highly interdependent relationship between the factors (Lessiter *et al.*, 2001; Wirth *et al.*, 2007), or as evidence that the processes are causally linked (Wirth et al., 2007). Here, factor analysis is not being used to create theory, but rather to illustrate something about the interactivity of factors. Second, even exploratory factor analysis can provide important evidence for one of the most important questions in presence research: Is presence unidimensional or multidimensional (Slater, 2003a)? An exploratory factor analysis can resolve this issue because it allows one to partial out items from a questionnaire which can then be used to form a subscale which can be subsequently validated. So, for example, if a factor analysis reveals some new factor, one can set up an experiment where one manipulates immersion to see its effect on this factor – if it is related to presence, one can expect differences between immersion conditions. One could still argue (as suggested by Slater, 2003a) that this happens not because that factor is a part of presence, but because it is caused by presence; but this then is an issue not of presence theory, but of presence definition; and no amount of data can resolve definition debates (those must be resolved by consensus). In some sense then, it is true that factor analysis cannot generate theory. But for the problems currently facing presence, it is a very useful tool.

2.4.1 Self-report measures

Although it is common to refer to these measures as *subjective measures of presence* (for example in Schuemie *et al.*, 2001) perhaps to contrast them to physiological measures (which are referred to as objective measures, for example in Meehan, 2001), this review will not follow this tradition, as the word *subjective* has an erroneously negative connotation in the sciences. To avoid passing judgment on these measures on any grounds other than their validity and reliability (which can be evaluated objectively, in the scientific meaning of the word), this review will use the more accurate and descriptive term *self-report measure*.

These instruments have subjects give scaled reports of their experiences, typically using Likert or semantic differential items. These scales can be evaluated using psychometric techniques which can give good estimates of validity and reliability (Anastasi & Urbina, 1996). A large number of self-report scales have been developed, reflecting a lack of standardization in measurement in the field. This lack of standardization should itself not be a concern, as excursions into alternative measures are desirable in any developing field; The concern lies in the wide range of quality found in these measures. For example, Tromp *et al* (1998), for a single study, devised two items without any psychometric evaluation, while Lessiter *et al* (2001) used a four factor, 44 item scale which has been evaluated and refined with data from several thousand subjects in dozens of media conditions. This review will only focus on scales either with published psychometric properties, or those which have reached *de facto* status through wide use.

2.4.1.1 The Slater, Usoh and Steed (SUS) questionnaire (1994)

This early and highly influential questionnaire is perhaps the most widely used of all. It has not been psychometrically evaluated by its creators, but data from a moderate sample (n=101) showed a reasonable level of reliability (Cronbach's alpha of 0.77; Nunez, 2002). The questionnaire seems to be highly valid. Many studies using immersion manipulations show significant differences in the predicted direction (see Schuemie *et al.*, 2001 for a review of some of these studies). Although its validity has been criticized by its authors due to its low ability to distinguish between real and virtual environments in a single study (Usoh *et al.*, 2000), many studies show that it can distinguish between high and low immersion media conditions (Bracken & Skalski, 2006; Nunez & Blake, 2003b; Towell & Towell, 1997).

The SUS measures a single factor, spatial presence (the sense of being in the virtual space), with six items. Sample items from this factor include "I had a sense of 'being there' in the office space" (Not at all/Very much) and "I think of the office space as a place in a way similar to other places that I've been today..." (Not at all/Very much so). This factor is similar to that measured by the *Spatial presence* factor of the IPQ (see 2.4.1.3 below), the *Sense of physical space* factor of the ITC-SOPI (see 2.4.1.4 below), the *Spatial presence (self-location)* item of the MEC-SPQ (see 2.4.1.5 below), and the *Spatial presence* factor of the TPI (see 2.4.1.6 below). No doubt the reliability of the scale could be improved by increasing the number of items, but since the scale's publication, no moves have been made to update it.

There are two scoring methods for the SUS: One is the typical method of averaging the scores across all items, while the other method, which is generally only used by Slater and colleagues (for example Brogni *et al.*, 2003; Slater & Steed, 2000; Slater *et al.*, 1994; Usoh *et al.*, 2000). is to count the number of responses scoring 6 or 7, and then treat that number as a binomial distribution. This technique is intended to disambiguate between subjects who truly had a presence experience from the rest (Slater *et al.*, 1994), although it does not control for acquiescent subjects who have a bias towards giving high scores regardless of their experience.

2.4.1.2 The Presence Questionnaire (1998; 2005)

The first version of the PQ was not widely circulated, but the second version, published by Witmer and Singer (1998), reached fairly widespread use (for example, Darken *et al.*, 1999; Johns *et al.*, 2000; Mania & Chalmers, 2001; Schubert *et al.*, 2001). Recently, a third version has been published with a new factor structure (Witmer *et al.*, 2005). The well-known second version of the PQ contained 32 items, which were cluster analyzed into four factors (the term factor here is used loosely by the authors, not in the sense of a psychometric factor):

1. *Involvement/control*: How involved subjects become in the VE, and how much control they perceives they have. This factor is similar to the *Engagement* factor of the ITC-SOPI (see 2.4.1.4 below), the *Attention allocation* factor of the MEC-SPQ (see 2.4.1.5 below) and the *Engagement* factor of the TPI (see 2.4.1.6 below). Sample items are "How much did the visual aspects of the environment involve you?" and "Were you involved in the experimental task to the extent that you lost track of time?"

2. *Natural*: How natural the interactions with the environment are perceived to be, and how consistent the VE seems with reality. This factor is similar to the *Realness* factor of the IPQ (see 2.4.1.3 below), the *Natural* factor of the PQ (see 2.4.1.2 above), the *Naturalness* factor of the ITC-SOPI (see 2.4.1.4 below), and the *Perceptual Realism* factor of the TPI (see 2.4.1.6 below). Sample items are "Were you able to anticipate what would happen next in response to the actions that you performed?" and "How much did your experiences in the virtual environment seem consistent with your real world experiences?"

3. *Interface quality*: The subject's ability to concentrate on performing the main task in the VE, and how much the interface detracts from this. Sample items are "How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?" and "How much did the control devices interfere with the performance of assigned tasks or with other activities?"

4. *Resolution:* The ability to observe and manipulate objects from multiple viewpoints and distances. Sample items are "How closely were you able to examine objects?" and "How well could you examine objects from multiple viewpoints?"

The Cronbach's alpha for the PQ was estimated at 0.88 (n=152) by Witmer and Singer (1998), and another reliability analysis by Nunez (2002) found a value of 0.903, (n = 101). Immersion manipulation studies (for example, Mania & Chalmers, 2001; Nunez & Blake, 2003a; for example, Sallnäs, 1999; Youngblut & Perrin, 2002) suggest a satisfactory level of validity; furthermore, the PQ has been shown to moderately correlate with the SUS in two studies (r = 0.76 and 0.56; Nunez, 2002).
The PQ has been criticized by Slater (1999) on two grounds: First, the PQ measures the *subjects' estimations of theorized causes* of presence rather than presence itself. Second, the PQ conflates properties of the subject (such as their ability to focus attention) with properties of the system (such as the quality of the interface). While this does not reduce the criterion validity of the measure (its ability to predict the

effect of presence on other variables), it does reduce its construct validity and makes research with the scale considerably harder (Anastasi & Urbina, 1996).

The third version of the PQ was developed by taking the 24 items from the second version with the highest item-factor correlations, and adding eight new items (Witmer *et al.*, 2005). These 32 items were administered to 325 subjects taking part in seven separate studies, and then factor analyzed. The resulting four factors were (in decreasing eigenvalue order):

1. *Involvement:* The degree to which the medium captures subjects' attention and their involvement in the VE. This factor is similar to part of the *involvement/control* factor in the previous version of the PQ, the *Engagement* factor of the ITC-SOPI (see 2.4.1.4 below), the *Attention allocation* factor of the MEC-SPQ (see 2.4.1.5 below) and the *Engagement* factor of the TPI (see 2.4.1.6 below). Sample items are "How much were you able to control events?" and "How responsive was the environment to actions that you initiated (or performed)?"

2. *Visual fidelity:* This is the same factor as *resolution* in the second version of the PQ (it consists of the same two items).

3. *Adaptation/immersion:* How much subjects focus on the VE content and how quickly they adjust to the VE experience. This factor is similar to the *Involvement* factor of the IPQ (see 2.4.1.3 below). Sample items are "How much delay did you experience between your actions and expected outcomes?" and "How quickly did you adjust to the virtual environment experience?"

4. *Interface quality:* This is the same factor as second version of the measure.

This new version has a small improvement in reliability (alpha = 0.91), but by having been factor analyzed, allows for a more detailed interpretation of the results. It should be noted that unlike the second version of the PQ, the order of these factors is ranked by eigenvalue order (such that the first factor explains the most variance, followed by the second, etc.). Unfortunately, Slater's two major concerns (discussed above) have not been addressed, although the factor analysis eases the conflation between individual and systems factors somewhat.

The scale was designed for use in immersive or semi-immersive environments, and is not well suited for use in non-immersive media such as text (Nunez, 2002; Nunez & Blake, 2003b). Items such as *How natural was the mechanism which controlled movement through the environment* (item 6 from version 3 - Witmer *et al.*, 2005), *How well could you actively survey or search the environment using touch* (item 13 from version 3 - Witmer *et al.*, 2005) and *How easy was it to identify objects through physical interaction, like touching an object, walking over a surface, or bumping into a wall or object* (item 29 from version 3 - Witmer *et al.*, 2005) do not have much meaning for non-immersive media, and are likely to increase the error variance in the measure due to subject confusion. Unfortunately, such items are spread between the factors, so excluding a factor from the scale does not solve the problem.

2.4.1.3 The Igroup presence questionnaire (2001)

The IPQ (Schubert *et al.*, 2001), was constructed from 75 Likert-type items, some of which were developed for the scale, while others were taken from extant measures including the PQ (Witmer & Singer, 1998), the SUS (Slater *et al.*, 1994), Ellis *et al.*'s measure (1997 in Schubert *et al.*, 2001), Carlin *et al.*'s measure (1997 in Schubert *et al.*, 2001), Towell & Towell's measure (1997) and Regenbrecht *et al.*'s measure (1998 in Schubert *et al.*, 2001). During development, two sets of data were collected: one used all 74 items completed by 224 users of systems of varying degrees of immersion (Schubert *et al.*, 2001), and the other with but with a reduced set of items (immersion related items were excluded) on 269 subjects. These data were factor analyzed to extract three factors:

1. *Spatial presence:* The sense of being in a coherent virtual space. This factor is similar to that measured by the SUS (see 2.4.1.1 above), the *Sense of physical space* factor of the ITC-SOPI (see 2.4.1.4 below) the *Spatial presence (self-location)* item of the MEC-SPQ (see 2.4.1.5 below), and the *Spatial presence* factor of the TPI (see 2.4.1.6 below).

Sample items are “In the computer generated world I had a sense of ‘being there’” and “I felt present in the virtual space.”

2. *Involvement:* A sense of being captivated by, and of focusing attention on the VE. This factor is similar to the *Involvement* and *Adaptation/immersion* factors of the PQ (see 2.4.1.2 above), the *Engagement* factor of the ITC-SOPI (see 2.4.1.4 below) and the *Attention allocation* factor of the MEC-SPQ (see 2.4.1.5 below). Sample items are “I had a sense of acting in the virtual space, rather than operating something from outside.” and “I felt like I was just perceiving pictures.” (reversed item).

3. *Realness:* How real the VE seems, and how consistent the experience in the VE is with their experiences in the real world. This factor is highly similar to the *Natural* factor of the PQ (see 2.4.1.2 above), the *Naturalness* factor of the ITC-SOPI (see 2.4.1.4 below), and the *Perceptual Realism* factor of the TPI (see 2.4.1.6 below). Sample items are “How real did the virtual world seem to you?” and “The virtual world seemed more realistic than the real world.”

The current IPQ contains 29 items, and is available in German, Dutch and English; however, only the German version (the first developed) has been psychometrically evaluated, leaving the psychometric properties of the other two versions unknown (igroup.org, 2004). The reliability of the IPQ (from development data - Schubert *et al.*, 2001), is adequate, with Cronbach’s alpha scores in the low to mid range (around 0.70). As two data sets were collected during development, it was possible to conduct confirmatory factor analysis using the second data set on the factors extracted from the first (Schubert *et al.*, 2001). The similarity of factor structure and factor loadings indicates the factors are stable. There is less evidence of the scale’s validity. Although the IPQ is freely available, it has seen limited use. Brown *et al* (2003) found differences in the expected direction with an immersion manipulation. Similarly, Waterworth & Waterworth (2003) found higher IPQ scores for subjects experiencing concrete stimuli than abstract stimuli. Finally, Schulte-Pelkum *et al*. (2005) found that the sense of self-induced vection in an audio environment correlated positively with

IPQ scores. The scale was not designed for immersive systems, and most of the development data was collected using mostly desktop VR systems, and a few HMD and cave users (Schubert *et al.*, 2001). It is not known how the scale fares with non-immersive media (books, etc).

2.4.1.4 The Independent Television Commission's Sense of Presence Inventory (2001)

The ITC-SOPI was designed to be a cross-media measure, and has been thoroughly psychometrically evaluated (Lessiter *et al.*, 2001). Development began with a literature review which identified thirteen constructs of importance, and 63 Likert-type items were derived from these (Lessiter *et al.*, 2001). Data from six different immersion conditions (n=604) was factor analyzed to extract a four factor structure, which explains slightly more than a third of the total variance (Lessiter *et al.*, 2001). The four factors are (in decreasing eigenvalue order):

1. *Sense of physical space:* A sense of being in the VE space, and that objects and characters occupy the space. This factor is similar to that measured by the SUS (see 2.4.1.1 above), the *Spatial presence (self-location)* item of the MEC-SPQ (see 2.4.1.5 below), and the *Spatial Presence* factor of the TPI (see 2.4.1.6 below). Sample items are "I had a sense of being in the scenes displayed" and "I felt that all my senses were stimulated at the same time."

2. *Engagement:* A sense of psychological involvement with and enjoyment of the VE content. This factor is similar to the *Involvement* and *Adaptation/immersion* factors of the PQ (see 2.4.1.2 above), the *involvement* factor of the IPQ (see 2.4.1.3 above), the *Attention allocation* factor of the MEC-SPQ (see 2.4.1.5 below) and the *Engagement* factor of the TPI (see 2.4.1.6 below). Sample items are "I had a sense that I had returned from a journey" and "My experience was intense."

3. *Naturalness (Ecological validity):* The sense that the VE and its content are lifelike or realistic. This factor is highly similar to the *Realness*

factor of the PQ (see 2.4.1.2 above), *Natural* factor of the PQ (see 2.4.1.2 above), and the *Perceptual Realism* factor of the TPI (see 2.4.1.6 below). Sample items are “The content seemed believable to me.” and “I had a strong sense that the characters and objects were solid.”

4. *Negative effects:* Negative physiological effects (such as dizziness and eyestrain) – this factor is negatively correlated with the other three factors. Sample items are “I felt tired.” and “I felt nauseous.”

The final form of the scale retained 44 items over the four factors (physical space: 19 items; engagement: 13 items; naturalness: 5 items; negative effects: 6 items). The four factors are conceptually independent, so no single presence score can be computed; rather, each administration produces four independent scores which measure separate aspects of the experience (Lessiter *et al.*, 2001; Nunez & Blake, 2006).

The ITC-SOPI factor structure was verified by factor analyzing two random sub-samples of the original dataset. Cronbach’s alpha for all factors were acceptable (physical space = 0.94, engagement = 0.89, naturalness = 0.76, negative effects = 0.77). The scale also has a good degree of validity, having being used in a number of published studies. In most cases, the scale is sensitive to immersion manipulations in the expected directions (Dillon *et al.*, 2001; Freeman *et al.*, 2004; Lessiter *et al.*, 2001). It is also sensitive to variation in VE content (Dillon *et al.*, 2001). The large sample used to develop the scale, and the large number of published studies using the it make this arguably the best psychometrically understood measure available.

2.4.1.5 The MEC Spatial Presence Questionnaire (2004)

The MEC-SPQ (Vorderer *et al.*, 2004), is unique in that it includes three forms: long (eight items per subscale), medium (six items per subscale) and short (four items per subscale). The scale consists of Likert-type items measuring eight constructs derived from the MEC model of presence (see 3.3.4 in chapter 3):

1. *Attention allocation:* How much attention the subjects devote to the VE. This factor is similar to the *Involvement* factor of the IPQ (see 2.4.1.3 above), the *Involvement* and *Adaptation/immersion* factors of the PQ (see

2.4.1.2 above), the *Engagement* factor of the ITC-SOPI (see 2.4.1.4 above), and the *Engagement* factor of the TPI (see 2.4.1.6 below). Sample items are “I devoted my whole attention to the [medium].” and “The [medium] captured my senses.”

2. *Spatial situation model:* The subject’s confidence in the accuracy of their SSM (see 3.3.4.1 in chapter 3 for a description). Sample items are “I was able to imagine the arrangement of the spaces presented in the [medium] very well.” And “Even now, I still have a concrete mental image of the spatial environment.”

3. *Spatial presence (self location):* The sense of having been transported to the VE. This factor is similar to that measured by the SUS (see 2.4.1.1 above), the *Spatial presence* factor of the IPQ (see 2.4.1.3 above), the *Sense of physical space* factor of the ITC-SOPI (see 2.4.1.4 above), and the *Spatial Presence* factor of the TPI (see 2.4.1.6 below). Sample items are “I had the feeling that I was in the middle of the action rather than merely observing.” and “It was as though my true location had shifted into the environment in the presentation.”

4. *Spatial presence (possible actions):* The sense of being able to interact with and move around the VE. Sample items are “I had the impression that I could be active in the environment of the presentation.” and “I had the impression that I could reach for the objects in the presentation

5. *Higher cognitive involvement:* The degree to which the subject thought more more deeply about the VE content (i.e. beyond spatial properties). Sample items are “I imagined precisely what it must be like to further explore the world presented in the [medium].” and “The [medium] presentation activated my thinking.”

6. *Suspension of disbelief:* The degree to which subjects take a critical view of the experience and allow themselves to influenced by errors and inconsistencies in the VE. Sample items are “I concentrated on whether

there were any inconsistencies in the [medium].” (reversed item) and “It was not important for me whether the [medium] contained errors or contradictions.”

7. *Domain specific interest:* The degree of interested in and knowledge of the VE content. Sample items are “I have felt a strong affinity to the theme of the [medium] for a long time.” and “Things like the ones in the [medium] have often attracted my attention in the past.”

8. *Visuospatial imagery:* The subject’s inherent ability to form mental images and manipulate visual data mentally. Sample items are “When someone shows me a blueprint, I am able to imagine the space easily.” and “I can vividly imagine how small I would seem at the foot of a high mountain.”

Note that these constructs are not produced by factor analysis, but reflect the structure of the MEC model (giving the scale a high degree of construct validity in terms of that model). The constructs likely correlate, and the order in which they are presented above does not reflect any relative degree of predictive power. All three forms have been psychometrically evaluated (Vorderer *et al.*, 2004). The scale is available in English, German, Portuguese and Finnish, making it the most translated scale currently available.

The MEC-SPQ measures only experience and not system variables, as it aims to be used in various immersion conditions (Vorderer *et al.*, 2004; Wirth *et al.*, 2007). It is thus mmune to Slater’s criticism of the PQ (Slater, 1999). Although it combines measures related to the specific experience (such as *spatial presence* and *suspension of disbelief*) with inherent properties of the subject (such as *domain specific interest* and *visuospatial imagery*), it avoids conflation by dividing these into subscales, so that independent investigations can be made on their impact.

Due to its tight coupling with the MEC model, the scale has seen limited use, but has shown high reliability, with Cronbach’s alpha values (even for the short, 4-item form) around 0.8 in studies involving various media forms (Vorderer *et al.*, 2004). In terms of validity, the scale does well – it is sensitive to media manipulations in the appropriate direction (see for example Böcking *et al.*, 2004; Gysbers *et al.*, 2004; Laarni *et al.*, 2004; Vorderer *et al.*, 2004).

2.4.1.6 The Temple Presence Inventory (2000-2004)

The TPI is still under development by Lombard & Ditton. To develop the questionnaire, two conditions were used (n=600) – a high immersion condition (to maximize presence), and low immersion condition (Lombard *et al.*, 2000a). Subjects completed 103 items which were developed from five concepts of presence found in the literature (Lombard *et al.*, 2000a). This was factor analyzed into eight factors, and the 42 items highest loading factors were retained. The scale is interesting as it combines individual and social presence in one instrument. This review shall only consider the three factors which measure aspects of individual presence (in decreasing eigenvalue order):

1. *Spatial presence:* The sense of feeling in the VE, and that the others shared the space. This factor is similar to that measured by the SUS (see 2.4.1.1 above), the *Sense of physical space* factor of the ITC-SOPI (see 2.4.1.4 above) and the *Spatial presence (self-location)* item of the MEC-SPQ (see 2.4.1.5 above). Sample items are "How much did it seem as if the objects and people you saw/heard had come to the place you were?" and "To what extent did you experience a sense of being there inside the environment you saw/heard?"

2. *Engagement:* Psychological and sensory involvement with VE content. This factor is similar to part of the *involvement* factor of the PQ (see 2.4.1.2 above), the *Engagement* factor of the ITC-SOPI (see 2.4.1.4 above), and the *Attention allocation* factor of the MEC-SQP (see 2.4.1.5 above). Sample items are "To what extent did you feel mentally immersed in the experience?" and "To what extent did you experience a sensation of reality?"

3. *Perceptual realism:* The degree to which the VE matches real experiences, in terms of perception rather than content. This factor is highly similar to the *Realness* factor of the IPQ (see 2.4.1.3 above), the *Natural* factor of the PQ (see 2.4.1.2 above), and the *Naturalness* factor of the ITC-SOPI (see 2.4.1.4

> above). Sample items are "Overall how much did touching the things and people in the environment you saw/heard feel like it would if you had experienced them directly?" and "Overall, how much did the things and people in the environment you saw/heard look they would if you had experience them directly?"

These factors are similar to the ITC-SOPI, and although no correlation studies have been conducted, it seems likely that the TPI and ITC-SOPI would match up factor by factor. The TPI is highly reliable, with Cronbach's alpha scores between 0.91 and 0.79. Although not many studies have made use of the scale, the procedure used in its construction (comparing high-immersion to low-immersion conditions as well as a strong basis in the literature) promises a degree of construct validation.

2.4.2 Behavioural observation

Some researchers observe behaviours to estimate presence. If the behaviours are consistent with the demands of the VE rather than the real world, then this is taken as an indication of presence (Freeman *et al.*, 2000; Held & Durlach, 1992). Subjects are normally observed in relation to some specific element of the VE (such as interacting with another character), and presence is inferred from this. It is therefore an instantaneous measure of presence (i.e. presence is evaluated at particular points rather than over the whole experience).

Exactly what is observed varies between researchers. No one theory drives these measures, apart from the general principle that a subject acting 'as if' they are in the VE is likely to be experiencing presence (Freeman *et al.*, 2000; Held & Durlach, 1992; Slater, 2003b); hence this notion is also called 'behavioural realism' (Freeman *et al.*, 2000). Some of these measures have been shown to correlate with self-reports, suggesting that they have some validity (see for instance Meehan, 2001; Slater *et al.*, 1995c). Determining reliability is difficult, as each measure is generally designed and used only for a specific situation. Generally speaking, three classes of behaviour have been used as presence estimators: automatic social behaviours, reflex actions and posture/sway measures.

2.4.2.1 Automatic social behaviours:

Sheridan (1992a) and IJsselsteijn *et al*. (2000) suggested that the degree to which subjects engage in customary social gestures during a VE experience (such as reaching for a handshake, making eye contact, or engaging in conversational turn-taking) could be used as an indicator of presence. Bailenson and colleagues examined the interpersonal distance kept between the subject and an animated agent during a VE experience, and found that the average interpersonal distance was more like the distance kept between real individuals when characters made eye-contact with the subject (Bailenson *et al.*, 2001). This finding was later replicated and validated by the fact that some subjects believed the characters to be controlled by other subjects (Bailenson *et al.*, 2003).

Another behavioural presence indicator is emotional response. Huang and Alessi (1999) argued that facial expressions in response to emotional environments could be used as presence indicators. Ravaja *et al.*(2004), found support for this, by noting that particular facial muscles activate in response to specific VE content (see 2.4.3 below), but due to the limits of that study, this technique should not yet be considered as validated.

2.4.2.2 Reflex actions

Among the first suggestions for measuring presence by observation are those of Held and Durlach (1992) and Loomis (1992). Held and Durlach argue that a subject's response to unexpected VE elements can be indicators of presence. So, if an unexpected event occurs in a VE, a present subject should be surprised by it, and exhibit a startle response (Held & Durlach, 1992). This method was validated by Wilson *et al*. (1997) who found that subjects did indeed show startle responses to unexpected VE content. However, this simple idea is complicated by evidence that the startle response may have a more complex relationship with presence. A break in presence is a form of startle response (Slater & Steed, 2000) made in response to an unexpected glitch or inconsistent element in the VE; but in this case, contrary to Held and Durlach's expectations, the startle response is an indicator of no presence (Brogni *et al.*, 2003). Loomis (1992) proposed a more sophisticated version of this idea. He argues that behaviour can occur in response either to the elements of the VE (*distal attributions*), or to events on the display, such as glitches (*proximal attributions*).

Presence would only be indicated by responses which have distal attributions (the startle response, for example, can be made in response to both proximal and distal attributions, so does not fit the bill). Reflex actions in response to VE elements (such as dodging from away from objects in the VE) constitute such a class of behaviours (Loomis, 1992). This idea was evaluated by Slater *et al*. (1995c), in an ingenious study in which a radio was placed both the real and virtual environments. The distance between the was then manipulated as a tone was played from the real radio (Slater *et al.*, 1995c). Subjects were tasked with pointing to the radio which was playing the sound, and it was predicted that present subjects would point to the virtual radio rather than the real one. This method was validated by a significant correlation between the angle-error on the pointing task and self-reported presence (Slater *et al.*, 1995c).

2.4.2.3 Posture and sway:

Several studies have used how subjects respond to perceived motion by swaying to compensate for virtual motion (which would be a distal attribution – Loomis, 1992). Illusory self-rotation (vection) was one of the first forms to be considered. Cohn *et al.* (1996) found that subjects who were asked to reach for objects in a VE rotating on the yaw axis compensated for virtual rotation in proportion to the speed of rotation. A more complex design was employed by Ohmi (1998), who crossed degree of rotation with display type. More immersive displays led to more sway, and sway was synchronized to the VE under all display conditions (Ohmi, 1998). Freeman *et al*. (2000) and IJsselsteijn *et al*. (2001) used driver's perspective footage of driving along a winding track displayed either using a monoscopic or stereoscopic display. Subjects swayed synchronized to the stimuli, and the stereoscopic condition produced more sway. Furthermore, the degree of sway was correlated with presence self-reports (Freeman *et al.*, 2000).

A more sophisticated body sway metric was used by Emoto *et al*. (2004). They reasoned that looking at scenes with a small field of view would affect the equilibrium system, as the information provided does not match experience, and the subject would try to compensate with their posture (Emoto *et al.*, 2004). Therefore, as the field of view tends towards a natural value, equilibrium should return to normal and result in less posture compensation and sway. They indeed found that a wider fields-of-view led to less sway. This finding, as with Slater & Steed's use of the startle response in breaks in presence (Slater & Steed, 2000), illustrates the difficulty of interpreting behavioural metrics as simple direct correlates of presence.

2.4.3 Physiological measures

Several physiological variables are considered correlates of presence, although as with behavioural observation, no standard yet exists. An early study by Jorgensen *et al.* (1997) noted that VEs can produce measurable changes in physiology (in particular heart rate and galvanic skin response), and soon after efforts were underway to estimate presence this way. The most successful examples of this method are associated with Meehan (Meehan, 2001; Meehan *et al.*, 2002; Meehan *et al.*, 2003). Subjects were placed into a virtual environment in which they must navigate around a deep pit in the floor. Not surprisingly, taking part in the task is a stressful experience (Meehan *et al.*, 2002). The VE was implemented in various immersion conditions, from a four-wall cave (Meehan, 2001) to an HMD based, body-tracked system with passive haptic feedback (Meehan *et al.*, 2002; Meehan *et al.*, 2003). The studies examined changes in mean heart rate, skin temperature and galvanic skin response, and were validated both by behaviour observation, and by the SUS scale. The results are fairly impressive: change in mean heart rate and galvanic skin response were significant predictors of SUS scores (Meehan, 2001; Meehan *et al.*, 2002). Change in mean heart rate was also sensitive to changes in simulation display update rate, albeit non-linearly (Meehan *et al.*, 2003). Others such as Wiederhold *et al.* (2002), Preston (1998) and Zimmons and Panter (2003) have found similar effects, while Dillon *et al.* (2001) found a difference in heart rate across content conditions (exciting content/calm content). Meehan concludes that such measures can be used to evaluate presence, but only for stressful environments, because one needs in the environment a situation which should raise heart rate and lower galvanic skin response for present subjects (Meehan *et al.*, 2003); in effect, something to allow a distal attribution. Slater (2002) has criticized this type of measure for requiring artifacts to be inserted into the VE (such as the pit); he argues that it is contrived and impractical for general purpose applications (Slater, 2002).

Other physiological variables have produced similarly limited results. A study by Ravaja *et al* (2004) used facial electromyography (EMG), galvanic skin response (GSR) and change in heart inter-beat intervals (IBI) as well as the ITC-SOPI self-report scale. EMG activity related to experiencing a positive mood state (i.e. increased *zygomaticus* and *orbicularis oculi* activity and reduced *corrugator supercilii* activity) predicted spatial presence. Similarly, spatial presence was predicted by GSR activity which indicates increased arousal. They found no effects on the IBI (which contradicts Meehan *et al.*, 2002). Although these findings are encouraging, they cannot be generalized due to confounds in the design. The specific EMG indicators, for instance, show positive mood, but this is hardly surprising as they occur at the achievement of a goal. Although this activity is correlated with presence, it is likely that more present subjects became more involved in the game and thus felt more satisfaction (and hence positive affect) on completing the goal. It is unlikely (as conceded by Ravaja *et al.*, 2004) that this is a general purpose presence measure. GSR is more promising, but increased arousal is known to follow changes in focused attention. Although attention has been strongly associated with presence (Schubert *et al.*, 2001; Wirth *et al.*, 2007; Witmer *et al.*, 2005), such measures could only explain a small amount of presence variance, due to the large number of other factors involved.

2.4.4 Brain imaging

Brain imaging has recently been considered as a possible measure of presence. Unfortunately, both VR and brain imaging require bulky, sensitive equipment (both often relying on detection of magnetic fields), which may interfere with each other. Furthermore, brain imaging requires subjects to keep still to reduce measurement error due to muscular activity (Hoffman *et al.*, 2003). One study used used fiber-optic goggles to provide stereoscopic display to subjects inside an fMRI scanner (Hoffman *et al.*, 2003). Preliminary work showed the display could produce differing immersion levels measurable as changes in presence (Hoffman *et al.*, 2003), effectively opening the door to fairly immersive environments within the confines of the fMRI magnet.

EEG has given some interesting findings. Bischof and Boulanger (2003) placed subjects in a complex maze while monitoring theta band oscillations, which have been linked to hippocampal activity during navigation in rodents (O'Keefe & Reece, 1993). They found theta activity predicted specific behaviours in the maze (particularly making navigation errors). Although they did not measure presence, their results indicate that EEF measurement inside is VE is possible. Mikropolous *et al.* (2004) used EEG to examine decreases in alpha-band oscillations and increases in gamma-band oscillations, which are associated with shifts in attention and with visual awareness (Mikropoulos *et al.*, 2004). Measures were taken on seven subjects navigating four VE conditions, increasing in complexity and fidelity. As expected, EEG readings showed a decrease in alpha-band oscillations, and complementary increase on gamma-band oscillation. Mikropolous *et al*. conclude that because the EEG activity is sensitive to manipulations of scene fidelity, these wavelengths could be used as an objective measure for presence. However, this reasoning is flawed: First, complex environments (with more data to process) will lead to more attention focused on the VE, as habituation will take longer (Kosslyn & Thompson, 2003). Given the large manipulations used in the study (e.g. one condition was texture mapped while another was not - Mikropoulos *et al.*, 2004) the changes in EEG may only be indicating this phenomenon rather than presence; or, at best, the measure is only sensitive to broad-brush manipulations of immersion. Second, Mikropoulos *et al*. (2004) used only the EEG measure without independent validation measure, making the conclusion about presence essentially post-hoc.

2.4.5 Breaks in presence (BIPs)

This technique attempts to get around the problems of having subjects scale their own experiences (as is done in self-report measures). In a BIPS measure, subjects report on the occurrence events which are easy to detect, and because only their occurrence is reported on, problems of individual differences in relative scaling are overcome. Subjects are asked to report on “breaks in presence”, which is the sensation of a sudden change or shift in attention which occurs when one is suddenly moved away from being present in a VE. Before the experience, subjects are instructed to report each time they experience a “transition to reality”, which is defined for them as the awareness that they are in the laboratory where the experience is happening. Subjects report either by simply calling out “Now” (Slater & Steed, 2000), or by calling out loudly the reason for the break, for instance “Cable pull” (Brogni *et al.*, 2003). A large number of BIPs is associated with a low presence experience. For validation, Slater and Steed used a Markov chain model to predict presence from the distribution of BIPs over time (Slater & Steed, 2000), and the model was then validated using SUS data. Brogni *et al.* (2003) subsequently showed that a simple count of BIPs is a good predictor of SUS scores. The technique is therefore generally validated, although only by a few studies.

One problem with BIPs is the issue of false negatives. It is fairly obvious that a false positive cannot occur, because even if the subject did not truly experience a break and calls it, the calling of the break will itself induce a BIP. However, it is possible that a break could occur and the subject does not report it, and furthermore, the probability of a subject correctly reporting a break is probably inversely proportional to the difficulty of the experimental task. So, for example, if a subject is engaged in a difficult task, a BIP might go unreported as the subject allocates most resources to the task and forgets to call out the BIPs; this is a problem with all measures which give the subject a dual task (see Freeman *et al.*, 2000 for another example of a dual-task measure). Note however that this is a difficulty with the reliability and sensitivity of the measure, not with its validity.

2.4.6 Questionnaires versus other types of measures

All measures must serve two masters: reliability (measuring the construct with minimal error), and construct validity (measuring the actual construct it aims not, and no other) (Anastasi & Urbina, 1996). Showing that a scale is reliable is simple – well established techniques exist for doing so (Cronbach's alpha coefficient, factor analysis, test-retest techniques, etc.). Showing that a scale has construct validity is much harder (Anastasi & Urbina, 1996), and it is here that the measured debate has centered.

One side of the measurement debate opposes self-report scales on the basis that they are not able to measure presence at all (Slater, 2004; Usoh *et al.*, 2000). In an interesting study, Usoh *et al.* (2000) created a virtual environment of their laboratory, and placed 10 subjects into that VE and 10 into the actual laboratory with the same task, and then took SUS and PQ measurements. The prediction was that the real laboratory should produce more presence than the VE. The results were not as predicted. No significant difference between environments was found on either scale, although there was a significant difference in the expected direction when the number of high scorers (scoring 6 or 7) on the SUS was compared (Usoh *et al.*, 2000). Usoh *et al.* reached two conclusions: first, subjects will interpret questions to make sense in the given context, and respond on that basis. Second, presence questionnaires are not suitable for cross-environment comparisons due to their lack of sensitivity.

The study, although interesting, has some significant flaws. First, as discussed by Usoh *et al*, the power of the study is low due to the small sample. They argue that they expected a large effect, so a small sample should have been sufficient to show the effect. However, given that there is an indication of a difference (both from the significant difference between high scorers, and from the fact that the mean differences, although not significant, are in the predicted direction), it would have made sense for them to continue collecting data to see if the difference disappeared with a larger sample. The effect may thus simply be that it is smaller than expected, particularly given the fact that presence questionnaires had, at the time of that paper's publication, already shown a number of success at picking out differences between immersion conditions (e.g. Hendrix & Barfield, 1996a; Prothero & Hoffman, 1995; Slater *et al.*, 1995c). A second weakness in the study comes from the selection of criterion. While it is reasonable to expect subjects in the real world to feel more presence on average than subjects in a VE, it is not necessarily the case that self-report scales will work for real environments. Usoh *et al* argue that the questions from the questionnaires used make sense for real environments, but this proposition is hard to defend. Consider these items, form the perspective of a subject in the real world condition: "To what extent were there times during the experience when the office space was the reality for you?" and "During the time of the experience, did you often think to yourself that you were actually in the office space?" (from the SUS) or "Overall, how much did the you focus on using the display and control devices instead of the experience and experimental tasks?" and "How natural was the mechanism which controlled movement through the environment?" (from the PQ). Some subjects may have understood the intention of the item correctly, but many could have been confused and responded with an answer in the center of the scale to be safe. Regardless of what the subjects actually did do, it is safe to assume that this confusion translated into a degree of random answering, which would in turn increase the error variance, and make it harder to detect small differences. In this regard, it is telling that for both the SUS and PQ scores, the real world condition produced higher standard deviations, indicating more randomness (and possibly confusion) in response. The conclusion that subjects interpret questions in some context and respond on that basis is true (indeed, it is a fundamental principle of psychometrics – Anastasi & Urbina, 1996), but the significance of that conclusion is not that subjects are unable to report correctly on their experiences, but rather that there will always be a degree of idiosyncrasy to a subject's response. Some of it can be controlled for (such as by ensuring that the context in which they respond is meaningful and not confusing), but the rest must be dealt with by making use of large enough samples when making comparisons, to overcome the error variance.

The problems raised by Usoh *et al* are nonetheless worrying, because they do highlight some difficulties of working with questionnaire data. Another interesting paper from Slater (2004), presents the central problem in presence measurement: What is it exactly that we are measuring? Slater notes that a presence questionnaires should at the very least be able to distinguish between reality and a virtual environment, citing Usoh *et al* (2000). This is true in principle, but one cannot simply expect a questionnaire to work in contexts beyond those for which it was designed (if anything, Usoh *et al*'s argument is one against the particular questionnaires used, and not against questionnaires in general). A second objection raised by Slater (2004) is that presence cannot be measured by questionnaires, as these are suited to measuring mental states, and no presence researcher has yet identified the mental state which makes presence (Slater, 2004). While it is true that presence has not been defined as a unique mental state, it is not true that it does not exist as a well documented, independent phenomenon. Unlike "colorfulness", the straw-man concept ingeniously conjured up by Slater to show the dangers of reification, there is evidence that presence exists as a concept separate to the act of studying it. Several independent researchers (Freeman *et al.*, 2000; IJsselsteijn, 2004; Meehan, 2001) have reported that subjects placed in virtual environments respond to the stimuli presented to them as if they were real. This is an indication that presence indeed is a real phenomenon, which can presumably be measured somehow. Furthermore, researchers from other fields have independently identified and documented phenomena very similar to presence, such as placeness (Relph, 1976), Goffman frames (Rettie, 2004) and the transportation imagery model (Bracken, 2005a). Finally, Slater suggests that there are many ways to evaluate presence (including physiological measures), and it should not be the case that one method of evaluation should be used to the exclusion of others. This is of course entirely true, but somewhat misleading. Currently questionnaires are the preferred method of measurement because they are better understood than the alternatives. Although questionnaires are subjective and they do have flaws, there is a large body of literature which details how to develop, asses and refine both the reliability and validity of questionnaires (Anastasi & Urbina, 1996). A thoroughly evaluated questionnaire has a known degree of error, and a known degree of ability to correctly measure the construct of interest. Physiological measures however, while having the potential for enviably small degrees of error, currently have almost no validity at all – it is simply not clear what physiological variables are related to presence, or if there exists such a thing as an identifiable means of estimating presence using physiological or neural measures (although there is also no evidence suggesting that such a thing is *not* possible). Once physiological measures have been validated (a process which is slowly underway – see Brogni *et al.*, 2003; Friedman *et al.*, 2005), they may begin to see wider use.

2.5 The relationship between concepts and measures

Although there is debate around the relative value of various definitions and measures of presence, little attention has been focused on the relationship between concepts and measures of presence. For those engaged in empirical work, this relationship is critical, as it represents the link between research hypotheses and the interpretation of evidence (Turner & Roth, 2003). The definition one picks is that about which one wants to make a finding. The wide variety of presence definitions allows one to be quite specific about what the finding will describe. However, when the study is actually run, it is the measure chosen that defines what the effect is about: a choice of BIPs as the measure effectively means one is focusing on attention and perception related aspects of presence, while a choice of the IPQ effectively means one is focusing on spatial presence. Care must therefore be taken to find definitions and measures which are compatible. The biggest concern in such a situation is nullifying the construct validity of the measure which is used. In most cases, developers of measures go to great pains to collect and report evidence for the construct validity of their measures (see Lessiter *et al.*, 2001; Vorderer *et al.*, 2004 for examples of this process in the presence field), because a measure without construct validity measures some unknown construct and is therefore useless for drawing conclusions. In practical terms, this means that the choice of presence concept and measure are not independent. When engaged in empirical work, the choice of measure is particularly important, so it may be a good strategy to narrow the choice of presence definitions by considering only those which are implemented by measures of high validity and high reliability. Fortunately, as this review shows, the range of high quality presence measures is reasonably wide, and allows one some freedom to pick a measure not only in terms of its psychometric properties, but also of the specific presence concept required.

2.6 The concept and measure of presence used in this work

For reasons of expediency, this dissertation will deal only with individual forms of presence; that is, social forms such as co-presence and social presence will be excluded. The model which will be presented will be evaluated only for individual processing situations, and although it may hold true for social forms of presence, it has not been validated for those situations. The approach of this dissertation, as the following chapters will reveal, will be strongly cognitive. At the same time, this work is highly empirical, so only concepts which are associated with highly reliable and valid measures are practical choices. First and most important, this project requires a concept and measure which are able to capture surface processing (perceptual phenomena), as well as depth processing (semantic phenomena). This is necessary to ensure that the conclusions can be made about cognitive processes rather than simply perceptual processes. It should be made clear the distinction between "perceptual" and "semantic" effects made here. Although perception involves a large degree of semantic effects (such as semantic priming, etc – see Pinker, 1995), the distinction here refers to the origin of the data being processed cognitively. Perception is a process which begins with a sensory stimulus and ends with an activated concept in working memory (Plotkin, 1998); while semantic processing largely operates on information stored in semantic or conceptual form in declarative memory (Plotkin, 1998). Therefore, perceptual effects in presence are taken in this work to mean those effects which arise from the VE, and are subject to immersion effects, while semantic effects arise in the mind of the subject, and are thus subject to experiential, but not immersion effects.

Second, as suggested by Lee's evolutionary argument and Biocca's principle of evolutionary primacy, the concept chosen should allow one to make claims about various types of media (not simply immersive media), as presence is a phenomenon which pre-dates immersive technology (Biocca, 2003; K. M. Lee, 2004). Third, in order to remain as generally applicable as possible, the only concepts of presence to be considered will be the unifying principles; and finally, the concept must be capable of being operationalized for empirical work. Using these criteria, Lee's concept is excluded, as it does not lend itself well to operationalization (this is because Lee's proposal is closer to describing a mechanism for presence than a concept of presence). From the remaining two, Lombard and Ditton's definition (1998) is preferred over Nunez and Blake's (2001) because, although the latter is more directly specified in cognitive terms, the former has achieved more widespread use, and will therefore allow better comparison to existing research, and therefore a more solid foundation.

It should be noted that by 'presence' this dissertation specifically refers to *virtual presence*, that is, telepresence for the special case where the remote site is virtual (following Sheridan, 1992). Although it is reasonable to suggest that the findings here will apply to telepresence also, none of the supporting presented here studies use real remote sites, and for the sake of parsimony no such inferences will be made. However, in line with the evolutionary arguments presence by Biocca (2003) and Lee (2004), this dissertation considers the mechanisms which lead to virtual presence can also lead to presence in the real world, and therefore presence in the real environment is not considered as a unique case.

Choosing Lombard and Ditton's definition excludes a number of notable measures. Specifically, it excludes all physiological and brain imaging techniques, as there is no expectation of what changes one would see when a subject experiences mediation or the illusion of non-mediation. It is true that an approximation to this can be reached following Meehan's (2001) or Slater *et al.*'s (1999) approach of adding to the environment features which produce a known response under conditions of non-mediation, but this unduly trades off reliability for validity as discussed in 2.4.6 above. Of the remaining measures, the BIPs technique (Slater & Steed, 2000) is excluded as it operates only at the surface level (attention and perception), and gives no indication of deeper processing. The best remaining candidates are the PQ, the MEC-SPQ and the ITC-SOPI, as these have been through the most rigorous psychometric evaluations. Of these, the PQ is excluded, as it has had the least use with non-immersive environments. From the final two (MEC-SPQ and ITC-SOPI), the ITC-SOPI is preferred, due to its longer track record in research among various independent research groups (suggesting higher validity), as well as having had a larger sample during development, which undoubtedly increased its reliability (Anastasi & Urbina, 1996).

Chapter 3: A critical review of current significant models of presence

This chapter reviews the most significant current models of presence. Four major model families are identifiable in the literature: The two-pole and environment selection models associated with Slater; the three-pole model associated with Biocca; the focus-locus-sensus and layers of presence models associated with Waterworth, and the measures, effects, conditions model associated with Wirth and colleagues. Each of these will be examined using a standard template to allow direct comparisons between models. This template includes:

1. *Description of the model* – a summary of the model structure and history.

2. *Presence in the model* – a discussion of how the model views the state of presence, and how presence exists in the model structure.

3. *Summary of empirical evidence* – a critical summary of the most important empirical work supporting the model.

4. *Critical discussion of the model* – a critical examination of the model, including comparisons to other models, the evidence in favour of and opposed to the model. This discussion will consider the plausibility of the structures, and whether the model creates contradictions.

5. *How the model explains the five problems* – An examination of the model's power to explain the five problems for presence (these are detailed in 3.1 below).

A summary table of this review can be found in Appendix G. The models are presented in chronological order of publication, but due to ongoing development, there is considerable overlap. This review constrains itself to models of presence as a phenomenon, due to this dissertation's aim of evolving a cognitive model of presence.

This excludes models of how presence is related to other constructs, such as Zeltzer's Autonomy-Interaction-Presence model (Zeltzer, 1994) or the Immersion, Presence and Performance model (Bystrom *et al.*, 1999b).

3.1 Five problems for presence theory

The study of presence presents a few phenomena which are counterintuitive and thorny for models to explain. The literature contains five such problems of interest. For the purposes of this review, these problems will be presented as phenomena free from theory or explanation. The following sections define and describe each of the five in turn, and describe their importance to presence.

3.1.1 The book problem (Biocca, 2003)

A book stimulates only a single modality, with very low fidelity and with a low bandwidth stream of information. Furthermore, the reader has almost no control over their movement or interactions within the mediated space (Biocca, 2003). A book therefore represents an extremely low immersion system, and should, according to many current presence models, produce low presence experiences (Slater *et al.*, 1996). However, presence experiences while reading books are possible (Gerrig, 1993; Nunez & Blake, 2003b; Towell & Towell, 1997). A theory of presence must therefore be able to explain why this is possible (unless such a theory were to explicitly limit itself to being a theory of presence within highly immersive environments).

3.1.2 The physical reality problem (Biocca, 2003)

This problem can be considered the opposite of the book problem. The real world presents a continuous, high bandwidth, multimodal stream of information, and allows complete control over movement and interaction in the environment. Nevertheless, people sometimes experience no presence in the real world, due to daydreaming or being lost in thought (Biocca, 2003).

3.1.3 The dream state problem (Biocca, 2003)

The final three of Biocca's proposed problems is that presented by dreaming. Dreams can result in intense presence experiences, and yet the subject is receiving almost no stimuli from the real world, and there is no mediated environment involved in the experience either. Where then are they present? As with the book problem, one can argue that what Biocca calls presence in a dream is not presence at all, but again, one would have to define presence in a narrow, media specific form to exclude this type of experience.

Another interesting problem with discussing dreams in presence is that following Lombard & Ditton's (1997) definition of presence as the 'illusion of nonmediation', dreams should not be considered, as no mediation occurs. However, if one understands "nonmediation" in a more general sense to mean the case where *perception of an environment leads to a sense that nothing intercedes between the subject and the environment*, then dreams provide a particularly interesting category of experiences. During a dream, normal perception is circumvented, and rather than having the visual (or auditory, etc.) cortices stimulated via the sensory organs as during normal perception, they are stimulated directly by the reticular activating system (Hobson *et al.*, 2000). From this perspective, a dream is like the much vaunted science fiction situation where the VE display is done by direct stimulation of the brain. It is therefore interesting to consider the reticular activating system as a 'mediating system' during dreaming, as it may give hints of how advanced media systems may lead to presence experiences.

3.1.4 The virtual stimuli problem (Nunez, 2004a)

From a physiological point of view, a person only ever experiences one stream of external stimuli. For example, all light, regardless of whether it arises from a VE display or from the sun outside the laboratory window is received by the retina in the same way. All stimuli which encode a VE are converted into physical stimuli (light, vibrations in the air, etc) in order to reach the subject and for presence to occur (Nunez, 2004a). All stimuli, regardless of origin, are in fact real; they are all just energy arriving at the sense organs. Stimuli which arise from virtual sources (which can be called "virtual stimuli") are not tagged as being virtual and belonging to a special subset of stimuli. Virtual stimuli can share physical properties which will mark them as being different from the other stimuli. For instance, all stimuli related to the virtual environment might come from a small area of space, and may thus begin processing from a small set of adjacent retinal cells (Craver-Lemley & Reeves, 1992), or the virtual stimuli might be of a higher average intensity (Jonides, 1981). However, these properties (spatial location, relative intensity, etc.) must be inferred during perception, and can thus only be grouped together after they have been partly processed (Ungerleider & Haxby, 1994). It is therefore incorrect to state that there is a discrete number of environments or stimulus sets for a subject to choose from (such a "real environment" and "virtual environment"). There exists only one stream of stimuli from which a user can infer any number of environments. Because this process is inferential, top-down effects will play a major role (Wirth *et al.*, 2007). The problem of virtual stimuli is thus: How are certain stimuli recognized as encoding one coherent virtual environment, and how are stimuli outside of this set excluded from the presence experience?

3.1.5 The inverse presence problem (Timmins & Lombard, 2005)

Inverse presence was first described by Timmins and Lombard (2005). It occurs when real events are experienced as if they were mediated. This is most likely to happen when the events are unusual or emotionally intense, such as during the perception of great beauty or when being the victim of a crime. It is not clear how common inverse presence experiences are (the interview method used by Timmins and Lombard does not allow that inference), but the documented existence of 97 cases (Timmins & Lombard, 2005) suggests it needs to be considered by presence theory.

One can argue that inverse presence need not be explained by a theory of presence, as it is not experienced during mediation. Timmins and Lombard (2005) argue that inverse presence involves one class of experience (real) being confused for another (mediated), which is, according to Lombard and Ditton's (1998) definition of presence, the essence of the presence experience. From a psychological perspective, it is a very interesting phenomenon, as it indicates that the experience of mediation is not specifically tied to a particular class of stimuli, but can be freely associated with other sets of stimuli. Uncovering the conditions under which this sense of mediation is activated might shed light on how the converse (a feeling of non-mediation) occurs.

3.2 The five problems as a yardstick of model power

If one agrees that the five problems present important phenomena in presence, then one can judge the relative value of a presence model by how well it explains the five problems. In this review, each of the current major models will be compared in terms of their response to all five problems, to gain a comparative benchmark of their power, and to illustrate their strengths and shortcomings. This approach has been used in a limited way by Biocca, who used the book, physical reality and dream state problems as a measure of the increase of power of the three-pole over the two-pole model (Biocca, 2003), and by Waterworth who used the book and dream states problem to show the relative power of the focus-locus-sensus model of presence (Waterworth & Waterworth, 2001).

3.3 Current significant models in the presence literature

3.3.1 The two-pole / environment selection model

3.3.1.1 Description of the model

Although presented as a unified model, this is actually a composite review of a number of separate ideas, which do not formally exist as a model in the literature. The term ‘two-pole model’ was coined by Biocca (2003), and the term ‘environment selection’ was used by Slater and Steed (2000) to refer to the same family of concepts (this review prefers the latter term, as it better reflects the current sophistication of this model). This model is an evolution of the classic telepresence model, where an operator experiences being present at the remote worksite (Biocca, 2003). Sheridan (1992a) suggested that virtual presence could be understood simply as telepresence for the special case where the remote worksite is virtual rather than real. This idea was extremely persuasive, and led much of the research during the 1990s. This review will not focus on that early work, but rather on the more recent developments of that concept which have been informed by considerable empirical evidence.

Early versions – the “two-pole” model

In the “two-pole” model (a term coined by Biocca, 2003) subjects exist in one of two states (or poles) – either present in the virtual environment, or present in reality (see Figure 3.1 below). There is debate as to whether presence occurs by degrees (e.g. Wirth *et al.*, 2007; Witmer *et al.*, 2005), or whether it is binary (Slater, 2002) - see 2.3 in chapter 2. Some confusion surrounds Slater’s binary position, as his questionnaire (the SUS - Slater *et al.*, 1994), provides a continuous score. However, it should be noted that Slater’s practice when administering the SUS was to quantify scores such that only those scoring 6 or 7 would be considered as ‘present’, while the others would be considered as ‘not present’ (Slater *et al.*, 1994). To add to the confusion, Slater admits that continuous presence may indeed exist (Slater, 2002). Nonetheless, the two-pole model generally works for either binary or continuous concepts of presence, because at its core the model posits a simple one-dimensional progression between ‘present’ and ‘not present’ in the VE (Biocca, 2003). Generally speaking, ‘not present’ is not well defined, although having another system or environment which interferes with the VE of interest is central, as this model views presence as comparative (Slater, 2003a). The interesting questions arising are how a subject moves from the ‘not present’ to the ‘present’ pole (see Sadowski & Stanney, 2002), and what factors interrupt presence (Slater & Steed, 2000). A major aim of this research is to identify factors which affect and mediate presence, such that a ‘presence equation’ of factors and their relative contributions can be constructed (Kalawsky, 2000; Sas & O'Hare, 2001). These factors are categorized as being internal or external (Sadowski & Stanney, 2002); Internal factors are associated with the subject, for example culture (Fontaine, 1992), a tendency to become immersed (Witmer & Singer, 1998), or age and personality (Heeter, 1992). External factors (which are sometimes referred to as immersion factors - Slater *et al.*, 1995b) include display field of view (Hendrix & Barfield, 1996a), pictorial fidelity (Welch *et al.*, 1996), scene detail (Slater & Wilbur, 1997) and display resolution (Bracken & Skalski, 2006).

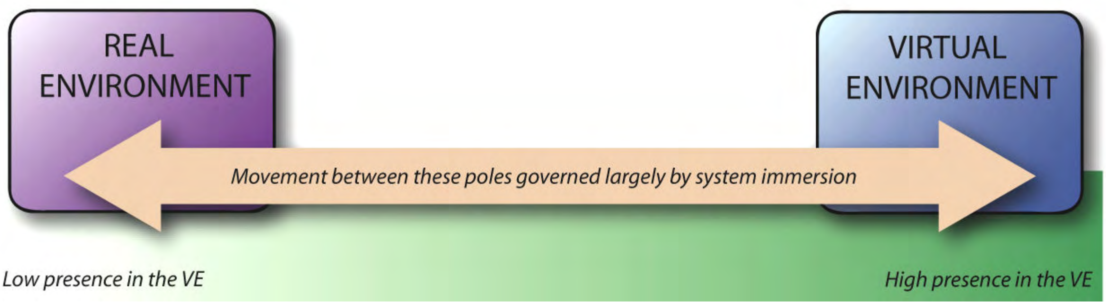


Figure 3.1: The two-pole model. The subject moves between the “real environment” pole (“not present”) to the “virtual environment” pole (“present”). The dynamic shift is determined largely by the degree of immersion of the system, although mediation by internal factors is thought to play a small role.

In many ways, the two-pole model (particularly the early versions) is not a model of presence as a phenomenon, but of presence as a desirable outcome for interactive systems. The dichotomy between ‘present in the virtual environment’ and ‘present somewhere else' (often appearing as 'present in the *real* environment') seems to develop naturally from this perspective – did the system manage to produce the desired effect (making users present) or not? And from that question, quite naturally, follow specific questions such as "what can be done to increase the likelihood of the desired result?" The current version of the two-pole model, which is better referred to as the *environment selection model*, is more sophisticated.

Current status - Environment selection theory

Environment selection theory assumes that subjects can only respond to and act in a single environment, even though they may be presented with several (Slater & Steed, 2000). This limit is imposed by their direction of gaze, limits of attention allocation, and other inherent factors (Slater & Steed, 2000). While a subject is present in an environment, it is perceived as a coherent whole (Slater and Steed use the term *Gestalt* to describe this coherence). To be present is thus to have selected one particular environment to respond to from among all competing environments (Slater & Steed, 2000). According to this model (outlined in Figure 3.2 below), a subject in a VE receives a continuous data stream from the VE, but also from other sources in the real environment (noises outside, temperature changes, etc) and from rendering errors in the VE system (Slater & Steed, 2000). Subjects in a VE thus always face two environments, and must choose one in which to be present. It is not clear whether only bottom-up data is able to force the selection of one environment over the other, or if top-down data has a role in this process (although this would seem a natural place for volitional processes such as 'the suspension of disbelief' – Slater & Usoh, 1993b).

This simple but convincing model has been subsequently refined by Slater (2002) by incorporating top-down and expectation based processing. As subjects interact in the VE, they form hypotheses about the VE (expectations for future data), and sensory inputs are tested under this expectation. If data which is consistent with the expectation arrives through other sensory channels (such as the addition of haptic feedback used by Meehan *et al.*, 2002), the VE becomes supported as a viable hypothesis. However, conflicting data (such as the tug of a cable or a rendering glitch), can cause the real environment to be selected, resulting in a break in presence (Slater, 2002).

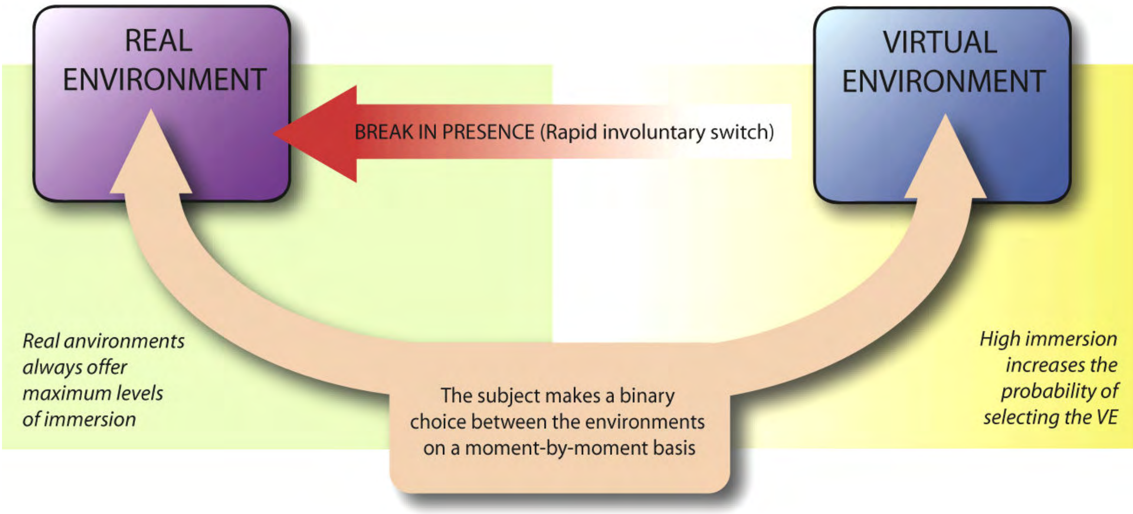


Figure 3.2: The environment selection model. The subject is either present in the real environment, or in the virtual environment. The probability of the subject selecting the VE depends largely on the level of immersion. Should a contradictory stimulus occur (such as a rendering artifact), the subject will rapidly switch back to being present in the real environment – a 'break in presence'.

The relationship between the amount conflicting data and the probability of a break in presence is not linear. Brogni *et al.* (2003) suggest that conflicting information could be attended out, incorporated into the current hypothesis, or lead to a break in presence. It is also not clear whether this will occur based only on the amount of conflicting sensory information, or if the content of the information has an effect. Depending on which environment which is selected, the subject will act differently; this allows the evaluation of presence by observing behaviour.

The process by which subjects switch between environments seems to be automatic, as There is no discussion of the possibility that a subject may willingly switch from one environment to another. In the reversible figure illusions which Slater uses to illustrate the Gestalt-like elements of the model (such as the young woman/old crone illusion and the duck/rabbit illusion – Slater, 2002; Slater & Steed, 2000) subjects can easily switch from one interpretation of the image to the other at will once they become aware that two alternatives are possible (Girgus *et al.*, 1977). However, this is probably not the case with presence. It seems that it is easier to stop being present in a

VE that it is to begin (although no empirical evidence to this effect exists). If this is the case, then it would be interesting to note the differences are between the reversible figure illusions and presence. It may be that, unlike reversible figures, there is little ambiguity in the environmental stimuli processed by the subject. The classic reversible figures are carefully constructed so that each line in the image is evidence supporting both hypotheses equally, and the small amount of data required to tip the decision in favour of one interpretation can be supplied by the subject willingly. In the presence situation however, it may be that the sensory stimuli tend to favour the "I am in the real world" hypothesis far more strongly than the alternative "I am in the VE" hypothesis; it may be impossible to provide the amount of top-down evidence required by will alone. This is because the suspension of disbelief argument assumes that all the relevant cognitive processes are mutable by the exertion of will; however, even for the case where a subject is able to exert a heavy degree of suspension of disbelief, the existence of automatic processing structures in the cognitive system (such as the automatic activation of the folk-psychology module – Baron-Cohen, 1995 – and the automatic association of active concepts in declarative memory – Atkinson & Shifrin 1968) may in fact prevent this from occurring, as some processes simply may not have any input from volition.

3.3.1.2 Presence in the model

Both forms of the model discussed above follow a consistent view of presence: Presence is the subjective sense that one is in the VE of interest, rather than in any other place (Slater *et al.*, 1994). Essentially, a subject exists on a continuum between the two poles of "virtual environment" or "real environment" (Biocca, 2003). The more recent reformulation by Slater and Steed (Slater & Steed, 2000) removes the emphasis of being present in either the virtual environment of interest or some other specific place, to a more generic sense that the subject is either present in the environment of interest, or in some other place, which is not defined. This improves the definition in some sense, as it simplifies operationalization; one need no longer worry about where in particular the subject felt present, in some sense obviating the need for a "third pole" as proposed by Biocca (2003) – see 3.3.2 below.

3.3.1.3 Summary of empirical evidence

Evidence for the two-pole model

The two-pole model simply predicts that particular factors (especially immersion factors) will increase the probability of the subject feeling present in the environment of interest rather than anywhere else. It is perhaps not so much a model as a loose framework upon which a program of raw empiricism has been based (see Kalawsky, 2000; Sas & O'Hare, 2001 for a discussion of the structure of this program). Much work has been done into uncovering factors important to presence, and much of it has been replicated. As discussed in 3.3.1.1 above, much of the impetus in this research is driven by practical concerns of creating “presence producing systems”.

This review will not consider studies which take into account internal factors (that is, subject related factors), for two reasons: First, there exists no psychological or physiological model of how the individual factors interact with immersion factors in the tradition of the two-pole model, and therefore no firm predictions related to internal factors (Biocca, 2002; Sheridan, 1992a). Second, studies which consider internal factors almost always have a more complex model to test, and are therefore are better discussed in the context of those models (see further discussions in 3.3.2, 3.3.3 and 3.3.4 below). Due to the large volume of studies in this area, this review will not go into detail, but rather categorize them by type of factor investigated.

Graphical display parameters

The largest body of work has considered graphical display parameters. Generally, these studies have a simple experimental design, where the factor of interest is manipulated, and its effect on presence is measured using a self-report scale (and in a few cases with behavioral observation or by counting breaks in presence). The general finding is that increasing the quality of the display by making its artifacts and limitations imperceptible increases presence (Lombard & Ditton, 1997). Some of the factors found to have an effect are:

- **Animation**. Animated scenes (where objects which would be animated in reality are animated in the VE) lead to more presence (Cho *et al.*, 2003)
- **Colour depth.** More colour depth leads to more presence (Barfield & Weghorst, 1993)

- **Display type**. Reality almost always out-performs any type of display (Allen & Singer, 2001; Hullfish, 1996; Mania & Chalmers, 2001) with a few rare exceptions (Usoh *et al.*, 2000). Multi-wall cave systems produce more presence than monitors (Axelsson *et al.*, 2001; Schroeder *et al.*, 2001). Head mounted displays generally produce more presence than monitors (Nichols *et al.*, 2000; Slater *et al.*, 1996; Slater *et al.*, 2000; Youngblut & Perrin, 2002), but this seems a weak effect, as numerous studies have failed to replicate the effect, perhaps due to the weight and discomfort associated with wearing such a display (Slater *et al.*, 1995a; Slater *et al.*, 1999; Youngblut & Perrin, 2002).
- **Display size**. Larger displays tend to lead to higher levels of presence (Lombard *et al.,* 2000a; IJsselsteijn *et al.*, 2001; Bracken & Botta, 2002; Bracken, 2005)
- **Display update rate.** Faster updates lead to more presence (Barfield *et al.*, 1998; Barfield & Hendrix, 1995; Meehan *et al.*, 2003; Snow, 1996). Note that this effect is likely non-linear, as Meehan (2002) found no difference between update rates of 10 and 15 Hertz.
- **Geometric field of view**. Wider displays lead to more presence (Hendrix & Barfield, 1996a; Prothero & Hoffman, 1995) although within limits - Allen and Singer (2001) showed maximal presence when using a natural FOV.
- **Level of detail**. More detailed, realistic scenes lead to more presence (Cho *et al.*, 2003; Shim & Kim, 2001; Slater *et al.*, 1995c; Welch *et al.*, 1996), although this effect may be weak, as other studies have failed to replicate it (Dinh *et al.*, 1999; Snow, 1996)
- **Resolution**. Higher resolution leads to more presence (Snow, 1996)
- **Stereopsis**. Stereo-enabled displays generally produce more presence than mono displays (Cho *et al.*, 2003; Hendrix & Barfield, 1996a; Snow, 1996)
- **Texture mapping**. Texture mapped scenes lead to more presence (Cho *et al.*, 2003; Snow, 1996)

Multimodality

A second clear finding is that systems which stimulate multiple modalities simultaneously lead to more presence that systems which stimulate a single modality. The following modalities have been examined:

- **Audio.** With very few exceptions, the addition of audio increases presence (Darken *et al.*, 1999; Dinh *et al.*, 1999; Hendrix & Barfield, 1996a, , 1996b; Nichols *et al.*, 2000; Sallnäs, 1999; Snow, 1996; Welch *et al.*, 1996).
- **Haptics.** Several studies have found that the addition of haptics or touch cues increases presence (Dinh *et al.*, 1999; Meehan, 2001; Sallnäs, 1999). However, a number of studies have failed to find an effect, suggesting that this is a weak factor (Insko, 2001; Lok *et al.*, 2003; Meehan, 2001)
- **Olfactory.** This factor has not received much attention (perhaps due to the engineering difficulties associated with implementing an olfactory renderer). The results are mixed – one study (Dinh *et al.*, 1999) found no increase in presence, but another (Hoffman *et al.*, 1999) found a small gain. Due to the small number of studies available, it is very difficult to draw a conclusion on this factor at this time.
- **Proprioception.** This modality can be implemented in immersive systems by the use of body tracking. This seems to be a strong effect, as it is replicated in almost all studies (Bystrom & Barfield, 1999; Hendrix & Barfield, 1996a; Snow, 1996)

System interface and interactivity

The two-pole model predicts that factors which provide cues to the subject that they are using a VR system could reduce or interrupt presence (Usoh *et al.*, 1999). Although VE interfaces are often examined as invariant system factors, the subject's proficiency with the interface will probably interact with the interface type. The following interface related factors have been examined:

- **Interactivity.** The more possibilities for interaction provided by the system, the more presence it generates (Snow, 1996); also, active roles in the VE lead to more presence than passive roles (Larsson *et al.*, 2001; Preston, 1998).
- **Movement.** Moving in the VE produces more presence than being stationary (Cho *et al.*, 2003); although this effect may be due to increased interactivity rather than increased navigation. The more natural the method of movement, the more presence reported by subjects; real walking generates more presence than passive motion or mouse control (Slater *et al.*, 1995a; Usoh *et al.*, 1999; Witmer & Singer, 1998).

Evidence for the environment selection model

The central tenet of this model is that the subject selects between environments to be present in; a corollary is that when a change in that selection occurs, it is experienced as a break in presence. Strictly speaking, there is no empirical evidence that subjects do select between environments, but there is evidence to show that certain distractions, particularly those associated with stimuli outside the VE, do lead to breaks in presence, which can be reported, and are associated with self-reports of presence. The first study to show this used SUS scores to predict the reported breaks in presence (Slater & Steed, 2000). Another similar study found a negative correlation between number of breaks in presence and SUS scores obtained during six separate immersive VE experiences (Brogni *et al.*, 2003). Finally, Vinayagamoorthy *et al.* (2004) also found a negative slope when regressing number of breaks in presence on presence questionnaire data. It is important to note that these studies do not show that environment selection takes place; it is evidence that the end of presence is a reportable experience. One can argue that the fact that during post-VE interviews some subjects did report experience a sensation of "switching" between environments supports the switch. However, these reports should be considered contaminated by the instructions given to subjects on how to report a break in presence (see 2.4.5 in chapter 2). Nevertheless, this evidence strongly supports the notion that the number of interruptions during the VE experience can inhibit the subject's presence.

3.3.1.4 Critical discussion of the model

The two-pole model has already been thoroughly discussed and criticized in the literature; in particular Biocca (2003) has outlined important weaknesses in that model (see 3.3.2.1 below). These criticisms revolves around the central assumption that a subject must either be present in the virtual environment of interest, or in some other environment. Biocca argues the one can be present in *no physical environment*, such as when one is lost in one's thoughts (Biocca, 2003). Furthermore, because the two-pole model emphasizes the immersion-presence relationship, it does not allow for subjects becoming present in non-immersive media such as books (Biocca, 2003). These two criticisms are correct, but they cannot overcome the fact that the two-pole model has more supporting evidence for its central notion than any other presence model currently available; it is almost impossible to argue against the immersion-presence relationship.

However, it is important to consider this evidence within its limits. None of the available evidence shows that immersion is either necessary or sufficient for presence; all it shows is that one path to presence (among an unknown number of paths) is through immersion. A case in point is the role of content factors in presence. Almost all the studies reviewed in 3.3.1.3 above (indeed, most of the studies reviewed in this chapter) ignore VE content as unimportant to presence. From a purely methodological point of view, this is correct, as content is usually held constant across conditions. But this is not the case when comparing *across* studies, which may have vastly different content. Slater has explicitly stated that content is not an important factor in presence (2003a), based on evidence such as that cited in 3.3.1.3 above. However, in order to make such a claim, a similar body of evidence would have to show that non-immersion factors (such as content) have no effect on presence. The lack of evidence for content effects simply reflects scarcity of studies, not lack of effects.

Although the environment selection model is derived from the two-pole model, it is different enough to warrant an examination in its own right. The central notion in this model (that subjects select between competing environments as Gestalts), is supported by analogy using the reversible figures illusion (Slater, 2002). These illusions work because each line in the figure can simultaneously support one of two interpretations (e.g. duck or rabbit), such that the figure is completely ambiguous (Slater, 2002; Slater & Steed, 2000). However, when placed in a VE in a laboratory. subjects need to deal with a vast array or sensory information of varying degrees of intensity which they must form into a Gestalt. The subjects must select relevant stimuli while attending out the rest, based on their significance and task demands (Nunez, 2004a; Wirth *et al.*, 2007). Unlike a reversible figure, such a situation has no finite set of alternatives. Each subject constructs their situation in terms of its importance to themselves at that moment. A more fair analogy would be to consider the reversible figure as a picture on being on a piece of paper in a room. Some observers might see the duck, some might see the rabbit; but some, who may not be paying attention to the task, may only see the piece of paper and the experimenter; others may only see the décor in the room, and so on. It is true that all of these subjects are selecting between alternatives; but it is not the case that the number of alternatives is bounded by the stimuli manipulated by the experimenter.

The environment selection model also fails to explain why it is far easier for subjects to be present in the real world than in a VE (see for example Usoh *et al.*, 1999). If there is a selection being made, what factors lead to one environment consistently being chosen over another? One possible answer is immersion factors. The real world has higher resolution, more detail and stimulates more senses than any virtual reality system current available; therefore, it is selected more often. This is plausible, but it obviates the very model it is being used to support – it simply returns to the two-pole model ("more immersion means more presence") but with the added constraint that presence is now binary (the virtual environment of interest is either selected or not). Given that there is no theoretical position that categorically states that presence is binary (Slater himself stating that it might be continuous – Slater, 2002), this seems an untenable theoretical position. Nevertheless, the basic notion that some selection is happening during presence is interesting, because of the elegant way in which it explains the role of attention in presence (Biocca, 2003; Waterworth & Waterworth, 2001; Wirth *et al.*, 2007), and because of the break in presence experience (Slater & Steed, 2000).

3.3.1.5 How the model explains the five problems

The book, physical reality and dream state problems were defined as a reaction to the two-pole model (Biocca, 2003), so one can expect that these phenomena will not be explained well. Biocca has already discussed these three problems with regards to the two-pole model in detail (see 3.1 above), but not with regards to the environment selection model.

The book problem

At first glance, the environment selection model seems a good candidate for explaining the book problem. When reading, a subject has competing environments to select from (the environment in the book and the real world), and the subject can choose to read the book or simply look at it. However, this model still relies on the basic notion that immersion is required for presence, even if there is some recognition that cognition mediates the process (Slater *et al.*, 1994). From this perspective, information presented in non-immersive and non-embodied forms is extremely unlikely to lead to presence (Slater, 2003a), and this effectively precludes books from producing presence experiences, although reading a book can still lead to psychological engagement and enjoyment (Slater, 2003a).

The physical reality problem

In this situation, the subject is not processing external stimuli – they are lost in their own thoughts or preoccupations (Biocca, 2003). Again, the environment selection model seems a likely candidate to deal with this phenomenon: does the subject select the environment in their imagination, or the real environment around them? Due to the model's focus on immersion (which concerns only external stimuli), it is again difficult to explain this phenomenon. The model is not well equipped to deal with this problem due to its lack of an explanation as to what happens when someone attends out external stimuli (Biocca, 2003). It seems clear that when a subject experiences an environment, they process it into mental representations, some of which are mental images (Kosslyn & Thompson, 2003). Could bringing such images willingly to mind not lead to a similar (if impoverished) sense of being in the space? The environment selection model cannot respond to such questions as it lacks a coherent notion of what partial or continuous presence is and how it arises.

The dream state problem

Biocca (2003) presents this problem as similar to the physical reality problem, because at its core is the issue of presence in imagined environments. During dreaming, all external sensory stimuli are attended out (or if not, they are generally incorporated into the dream), and replaced by internally generated stimuli, some of which emulate bottom-up information (Hobson *et al.*, 2000). Although this situation can be considered a high-immersion situation (sensory stimuli have been replaced by the "virtual environment" of the dream), generally only a few modalities are stimulated, and dreams often contain a number of logical and perceptual inconsistencies. This makes it difficult to explain with the two-pole model (as discussed by Biocca, 2003). However, the environment selection model is well capable of dealing with dreams. Notice that in a dream, there is in fact only one environment (the dream), as all competing environments have been attended out. Therefore, even with the low levels of immersion generally found in dreams, the model has no problem explaining how one can feel present – the dream *must* be selected as the environment to be present in as it is the only choice available. Of course, this assumes that certain minimum criteria are met, for instance that the dream is in fact about a place and the subject experiences the space from the perspective of someone occupying that space.

The virtual stimuli problem

These models are not able to deal with the virtual stimuli problem due to their central assumption that presence is the selection between competing environments. Recall that to perceive an environment, the subject must first construct that environment as a coherent cognitive entity by selecting a particular subset of stimuli from the undifferentiated mass of stimuli arriving at the senses (Nunez, 2004b). Due to the limitations of human cognition, only a small subset of stimuli can be processed (Baddeley, 1986). Proposing that subjects are able to simultaneously construct and maintain several environments to select from violates this principle. One may counter by arguing that the model is not cognitive, but rather descriptive; an outside observer can enumerate several possible environments which can be constructed from the available stimuli, and then interpret the subject's behaviour as a choice between those environments. This is a true, but it fails to consider that subjects construct environment from both bottom-up and top-down data (Nunez, 2004a; Slater, 2002; Wirth *et al.*, 2007). Therefore, it may not possible for an external observer to predict or even describe the environment which the subject is experiencing presence in. Evidence for the importance of this comes from Nowak *et al.* (2006), who found that the presence in violent games was mediated by the degree of *perceived* violence in the game. To an outside observer, a game has a constant degree of violence; however, due to individual differences, subjects may construct the environment as being more or less violent, which in turn affects their presence experience. The environment selection model could show that the subject is present in a violent game by examining their behaviour; but the detail required to differentiate between two subjects who perceive different degrees of violence and therefore have different presence experiences could not be achieved by this model.

The inverse presence problem

The inverse problem arises when a subject mistakes the real scene for a mediated one (Timmins & Lombard, 2005). The classic two-pole model is not able to explain this phenomenon, as it defines a strong distinction between "real" and "virtual" in terms of immersion. Presence arises (almost automatically) as a function of having sufficient immersion. The two-pole model makes the sensation of being in the real world the standard against which less immersive mediated experiences are compared. It is therefore almost impossible to understand how a completely immersive situation (such as the real world) could lead to a sense of less presence. The environment selection model fares little better. In this model, one environment is selected from competing environments by the subject to feel present in. However, in the situations of inverse-presence presented by Timmins and Lombard (2005), there is usually only one environment available. The environment selection model predicts that at worst, a low immersion environment could be selected for presence, and in such a case, the presence experience would be low; but in its current form it cannot explain why a subject should experience a highly immersive environment which they have selected as a low presence experience.

3.3.2 Three-pole model

3.3.2.1 Description of the model

This psychological model sees presence as moving in a space defined by three idealized poles (see Figure 3.3). These poles represent complete presence in a physical space, complete presence in a virtual space, and complete presence in a mental imagery space. The model contains no explicit notions of immersion or display technology. In fact, such concepts have been removed from the model for two reasons:

1. Biocca argues that the inclusion of system and immersion variables is not relevant to explaining psychological states such as presence. The idea that immersion leads to presence (the 'sensorimotor immersion assumption' in Biocca's terms) was dictated by engineering expediencies rather than psychological theory (Biocca, 2003). For researchers working with an engineering hammer, presence naturally seemed like an immersion nail. Biocca argues that a general model of presence for use in various media must reconsider the role of technology in presence, rather than assuming it as a necessary condition.

2. Presence must have existed before VEs, as the psychological mechanisms involved must be evolved (the 'evolutionary primacy' principle - Biocca, 2003). The certain media lead to presence is an indication that something in those media capitalizes on particular aspects of cognition (as perceptual illusions do - Slater, 2002). It then follows that explanations of presence should be independent of media, and conversely, that any medium could potentially induce presence. It should therefore be the psychological mechanisms involved in presence which should take center stage in a presence theory, not the display (Biocca, 2003).

The three-pole model is essentially an elaboration of the two-pole model. Each of the three poles (physical, virtual and mental imagery) represents stimulus sources which can lead to presence in that space. For example, attention focused on a display encoding a VE will lead to a high degree of presence. As with the environment selection model, having attention divided between poles leads to reduced presence. These stimulus sources dynamically change and possibly compete with each other, causing presence to be an oscillating phenomenon. Here an important difference exists between this model and the environment selection model: Biocca (2003) specifically allows the possibility that cues from the three sources could interact or be integrated into each other to form a coherent presence experience (as opposed to the environment selection model that sees the two stimulus sources always interfering with each other).

The relative contribution of each pole to presence is controlled by two cognitive processes:

1. *Spatial attention:* According environmental changes and task demands, attention will shift between the three poles during the experience. A loud noise, for instance, will demand attention to itself, which will change the relative contribution of the three poles; or a difficult spatial task may lead to attention being shifted towards mental imagery space during a portion of the experience.

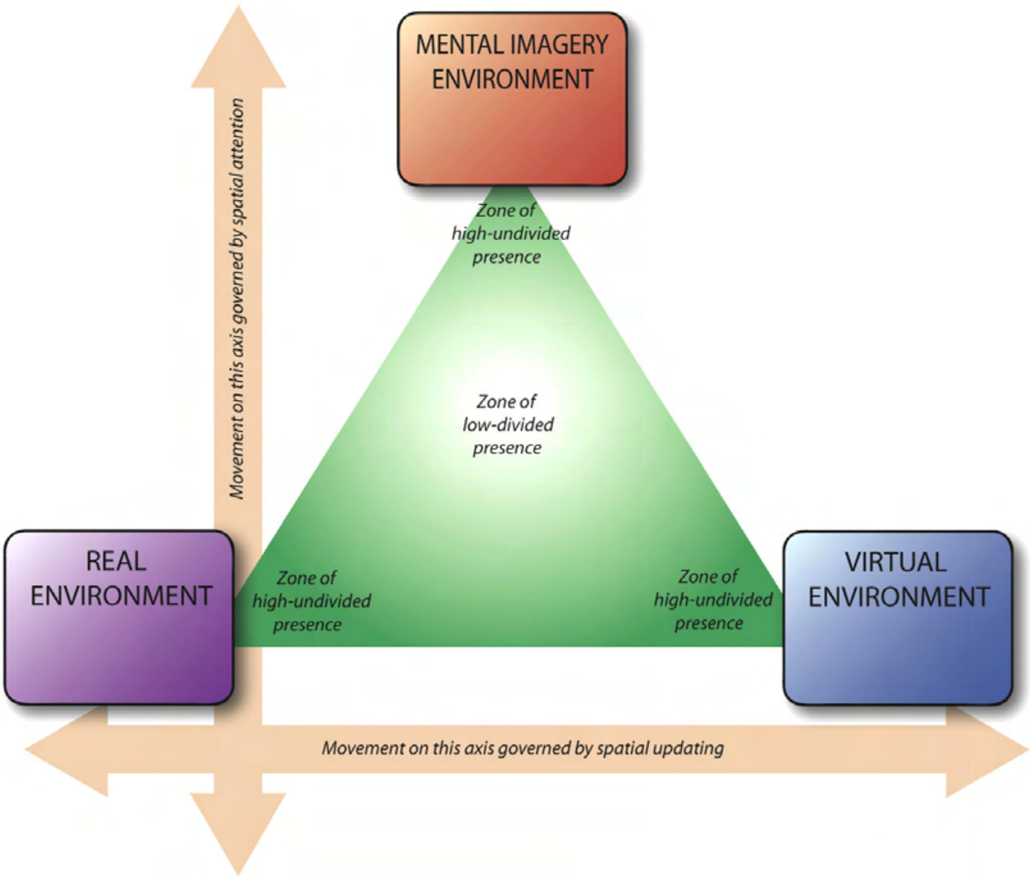


Figure 3.3: The three-pole model. The subject's experience is always located somewhere inside the green triangle; closer to the poles (real, virtual and mental imagery environments), the experience is *undivided presence*, while at the center is a *divided presence* experience. The experience moves dynamically in this space according to changes in spatial updating and spatial attention (tan arrows).

2. *Spatial updating:* As the experience progresses, the subject's model of the space and their relation to it is updated (Biocca, 2003). Such updates can lead to changes in the relative contribution of the three poles. This process is associated with bodily movement (either in terms of moving sensors within the space or affecting the space with the body), and as such is theorized to affect mostly shifts between the virtual and physical spaces.

3.3.2.2 Presence in the model

Presence is conceptualized as the processing of the cues from the three poles, with the weight of each pole being determined by spatial attention and spatial updating. Presence is thus a highly dynamic process, as the relative contributions change constantly, making presence continuous rather than binary. Subjects who have a high contribution from one of the poles experience ‘undivided presence’, while those who find themselves in environments where more than one pole is making a significant contribution would be said to be experiencing ‘divided presence’. The undivided form of presence is considered to be ‘high presence’, while the divided form is considered to be ‘low presence’. This model therefore considers focused, exclusive attention on one set of cues as a fundamental requirement to presence.

It is not clear how the transition happens between focusing on a stimulus set and presence in that set occurring. Although it is clear that focused attention is necessary, it is not obvious that it is sufficient. Strictly speaking, the three-pole model is not a model of presence *per se*, but rather a model of presence shifts. The model does not really need to explain how presence comes about in order to explain how subjects move from being present in a VE to being present in a real environment. If this is indeed the case, then presence exists as an ephemeral substance moved by the winds of spatial attention and spatial updating between the three poles.

3.3.2.3 Summary of empirical evidence

There has not been much empirical validation for this model. However, as it is an extension of the environment selection model, one can consider a certain subset of the evidence for that model as valid here. Specifically, all evidence which points to presence being a phenomenon which shifts continuously and dynamically between two cue sources can be applied, as can the subset of studies which finds evidence for the role of focused attention in presence (see the summary of the evidence for the environment selection model in 3.3.1.3 above for this work).

The question of whether mental imagery spaces can lead to presence, and how this can lead to presence in other types of spaces was the subject of a study by Baños *et al.* (2005). Subjects were placed into either a VR condition, where they experienced a park using a desktop VR system, or an imagery condition, where they were asked to imagine the park. Both groups were measured using the SUS at the beginning, middle and end of the experience. Averaged over thee measurements, there was no difference in presence between the VR and imagery condition, supporting the three-pole model’s concept of presence in mental imagery spaces. However, when considering how presence scores changed over time, then an effect became evident: the VR group scores increased over the experience, whereas the imagery group scores decreased (Baños *et al.*, 2005). This suggests that while the model is correct about the existence of the third pole, it may not be equivalent to the other poles in terms of maintaining focus.

There is some interesting indirect evidence for the importance of spatial updating on presence. Barfield and Hendrix (1995) had subjects perform a visual search task in a VE where the simulation update rate was manipulated. Subjects in the high refresh rate conditions reported more presence than those in the low refresh rate conditions. Spatial updating can account for this difference. As subjects move through the VE, they update their mental models of the space; however, if there is not timely feedback, then spatial updating would be inhibited, as the subject is placed in a moment of doubt as to the outcome of their input. Conversely, high update rates would give almost instantaneous feedback, allowing subjects to keep their mental models synchronized with the simulation. Another study which can be interpreted as evidence for the importance spatial updating is that by Slater *et al.* (1995c). In this study, subjects performed several spatial tasks either with or without shadows rendered in the VE. Subjects in the shadows condition reported higher presence than those without (Slater *et al.*, 1995c). A shadow provides additional information about the environment (such as position of light sources, geometry of the surface on which the shadow falls, etc.) and the motion of objects relative to each other (Kersten *et al.*, 1997). Shadows in the environment would therefore allow easier updating of the subject's mental model by providing more information about changes in the space.

3.3.2.4 Critical discussion of the model

The stated purpose of this model is to deal with three particular problems in presence theory (the book, physical reality and dream state problems), and it does this very well. If one accepts that these are important problems for presence, then one can claim that the three-pole model is a useful evolution of the environment selection model. This model can thus claim almost all evidence which supports the environment selection model, and simultaneously explain Biocca's three problems, which is an impressive feat. However, the model does have some weaknesses which are worth mentioning.

The first problem is shared with the two-pole model. Presence moves dynamically between three poles, such that the sum of the contributions of the poles is constant (or, stated differently, the three poles trade off against each other). This implies that there is a finite resource available for this process. However, the model does not state why such a limitation should exist, or what the limited resource is. One can infer that the limitation is tied to spatial attention and spatial updating. For instance, attention is known to be capacity limited (for example, see Baddeley, 1986). It would be useful if the model explicitly defined the limitation, as that would improve predictions of how spatial attention and spatial updating act to move the presence between the poles.

As most presence models do, the three-pole model gives primacy to the role of attention in presence. However, this model places particular emphasis on *spatial* attention, and the selection of stimuli in terms of their position in space. It is true that the spatial origin of a stimulus can change the focus of attention (Vecera & Rizzo, 2003), and that stimuli can be aggressively filtered out based on their spatial location (Hopf *et al.*, 2006). However, attention can shift due to a number of low level factors such as stimulus intensity (Lu & Itti, 2006) as well as high level factors such as priming (Maxfield, 1997). For example, breaks in presence, which are effectively rapid shifts between two of the poles, are often triggered by changes in attention which arise from non-spatial sources (such as rendering errors or inconsistencies in the simulation - Slater & Steed, 2000). It seems then that the role of non-spatial attention may have been underestimated by the model.

Another problem associated with emphasizing spatial attention lies in considering how presence in a physical space ends. The model defines the most intense presence experiences as occurring when subjects focus all their attention on a single pole As movement between the poles is partly determined by spatial attention, this implies that if a stimulus comes from an unexpected location, it will lead to presence being drawn away towards it. When a subject is using a VE, it is simple to define what an 'unexpected location' is (outside of the experimental room, for example - Slater & Steed, 2000). However, as physical space, by definition, occupies all space, there can be no notion of an unexpected location for a stimulus to occur from. Consider this thought experiment: I am sitting, present in real a forest. A laptop is running somewhere behind me without my being aware of it, which is showing a virtual office. Suddenly, a virtual telephone rings in the laptop – all I hear is a phone ring behind me. I am likely to think it is a real phone behind me, and on turning to look, be surprised to find it is a virtual phone. Now consider another situation: I am in the office, but watching a laptop showing a virtual forest. I hear a bird sing behind me. I am probably more likely to think that it is a real bird outside my office window than a virtual bird in the VE. Even though in both scenarios the stimulus comes from an unexpected location, and is of a type which is unexpected for the environment, it is likely easier for attention to shift away from the virtual space towards the real than vice-versa. This is partly because the primary expectation is that all stimuli come from the physical space; after all, attention (spatial or otherwise) exists is to direct cognition to some stimuli in the outside world in order to alert us to sudden changes in the environment which may represent danger or other interesting events (Sperber & Hirschfeld, 1999). The three-pole model however, posits that physical, virtual and mental imagery spaces as equally important in terms of attention, which may not be the case.

The three pole structure of the model also raises some questions. It is clear that there is a set of cues from physical space which can lead to presence; Also, there is some evidence that a set of mental imagery cues that can produce presence (for example Baños *et al.*, 2005). However, it is not entirely obvious that there exists such a distinct entity as a virtual space, comparable to either physical or imagery spaces. This returns to the problem of virtual stimuli. If a VE can only be processed after its existence has been inferred from stimuli originating in the physical space, then the physical space and the virtual space are not independent poles, as proposed by the model.

3.3.2.5 How the model explains the five problems

The book problem

This model was constructed as an explicit solution to the book, physical reality and dream state problems, so it should be no surprise that these three problems are dealt with rather well. The book problem arises from the fact that subjects have presence experiences from low sensorimotor immersion environments. The solution exists in the mental imagery pole. Reading a book involves many cognitive processes including the creation of mental models of the space described in the book. Spatial attention is drawn to these mental models, and spatial updating is performed on the models, which leads to presence being dawn towards the mental imagery pole. Provided spatial attention is not drawn away from the imagery space pole, and that the space is successfully updated, the subject can sustain their presence in the book. However, because reading involves a balance between attending to the mental imagery space created by the text, and to the visual task of reading, the presence experience will usually be in a zone of low-divided presence (Biocca, 2003). This explains why reading usually produces a less intense presence experience than more immersive media (Nunez & Blake, 2003b).

The physical reality problem

Essentially, this problem asks how subjects in the physical world (a high sensorimotor immersion experience) can sometimes not be present. Again, the solution comes from a shift from the physical space pole to the mental imagery pole. If spatial attention is shifted away from the physical space to the mental imagery space (by daydreaming, for instance), then presence in the physical space will diminish. In this model, all three poles are equivalent, so such a shift is consistent with the model. If some task is being performed in the physical space, then some attention will still be devoted to processing it, and the result will be a shift into a zone of low-divided presence (Biocca, 2003).

The dream state problem

The dream state problem is seen as an extreme form of the physical reality problem. In this case there is almost no input from physical or virtual sources, but a rich set of semi-random stimuli from activations sources in the subject's brain (Hobson *et al.*, 2000). During a dream, brain activation of the parietal regions (which subserve spatial cognition) can lead to a coherent mental imagery model of some space (Hobson *et al.*, 2000). If this occurs, then this model can attract spatial attention to itself, and could potentially support spatial updating, leading to presence (particularly if the special case of lucid dreaming turns out to be true - see LaBerge, 1980). Because there is almost no competing stimulation from any other sources, attention can be focused on the mental imagery space exclusively, and a high-undivided presence experience can occur.

The virtual stimuli problem

This model is not able to deal with the virtual stimuli problem, due to the existence of the virtual space pole. The essence of the virtual stimuli problem is that there is no external distinction between stimuli which originate from the virtual source of interest and from other physical sources. Stimuli can only be inferred to have originated from a virtual source once they have been partly processed (Nunez, 2004a). This model guards itself partly form this problem by positing that the movement of presence between the poles is partly due to spatial attention. A subject experiencing a desktop VR system knows which stimuli arise from the virtual space and which arise from the physical space partly by virtue of their location in space – the virtual space stimuli arise from a particular rectangular region (the screen), and all other stimuli are from the physical space. However, this alone does not solve the problem – if for example, one of the pixels of the monitor were to become stuck on a particular colour (say red), then the stimulus of the sole red pixel would probably not be experienced as originating from the virtual space, but from the physical space. Spatial information alone is thus not enough to explain the distinction. The virtual space pole is thus not equivalent to the physical space pole or the mental imagery space pole, as these two poles provide stimuli directly (without the need inference), and the movement from the virtual pole to either of the other two cannot easily be explained by reference to shifts in spatial attention and spatial updating.

The inverse presence problem

In inverse presence, a real world phenomenon is experienced as if it were mediated or virtual. One would have to imagine a situation where attention and spatial updating are firmly focused on extreme end of the physical space pole, and yet the experience is as if focus were on the virtual space pole. This is a difficult phenomenon for this model to explain, as the model does not separate between the perception of presence and the source of the stimuli. The strong link between the source of the stimuli and the type of experience makes it almost impossible to explain inverse presence from this perspective. Also, the low cognitive level of this model (which works mostly in terms of perception and attention) makes it difficult to make use of other concepts such as memory and expectation based processing (as suggested by Timmins & Lombard, 2005) to incorporate inverse presence without major revision.

3.3.3 The Focus, Locus, Sensus / Layers of presence models

3.3.3.1 Description of the model

This model was first proposed by Waterworth & Waterworth (2001), refined in 2003, and further developed by Riva, Waterworth and Waterworth (2004). The 2001/2003 form of the model, called the Focus, Locus and Sensus (FLS) model, does not explain presence itself, but rather explains the experience of being, how this shifts between real and virtual worlds, and how subjects move from awareness to non-awareness of the external world. It proposes three orthogonal dimensions (see Figure 3.4 below):

1. *Focus:* The two extremes of this dimension are *presence* and *absence*. Presence occurs when attention is focused on perceptual processing of the concrete world outside the self. If processing becomes abstract, then attention is turned inwards, and the subject becomes absent in the world (Waterworth & Waterworth, 2001). As attention is limited (Baddeley, 1986), there is an inherent and conscious trade-off between presence and absence along this dimension. A divided state is possible, and indeed occurs in most normal experiences (E. L. Waterworth & J. A. Waterworth, 2003; Waterworth & Waterworth, 2001).

2. *Locus:* This dimension has the extremes of 'real environment' and 'virtual environment', but more specifically describes whether the subject experiences the environment directly, or mediated in some way. At the 'real environment' pole, the subject is directly embodied in the environment, while at the 'virtual environment' pole, the subject experiences the environment through an interface or some other set of hermeneutic relations (Waterworth & Waterworth, 2001). This pole captures the essence of Lombard & Ditton's (1998) notion of presence as the illusion of non-mediation. At the 'real environment' pole, there is no mediation (or at least, no perception of mediation), while at the 'virtual environment' pole, the subject experiences obvious mediation.

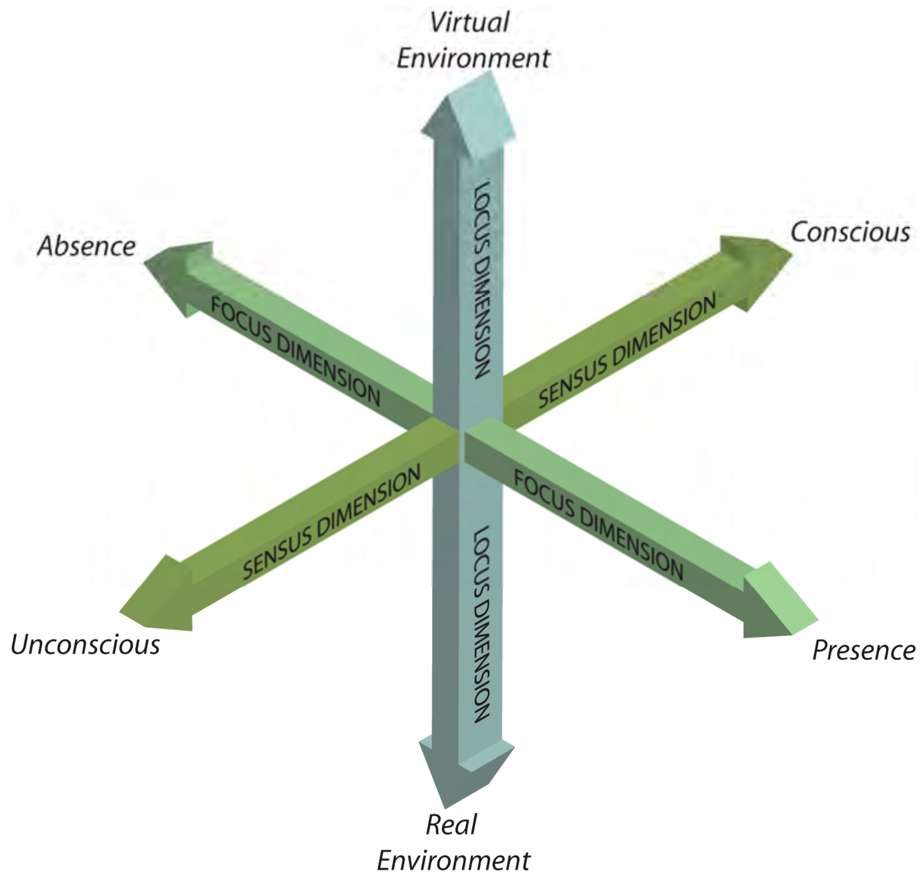


Figure 3.4: The Focus Locus Sensus (FLS) model. The subject's experience is a dynamically moving point in the space defined by the three dimensions.

3. *Sensus:* A novel aspect of this model is this dimension, which describes the subject's level of physiological arousal of the subject, between the extremes of *conscious* and *unconscious*. This dimension interacts with the focus dimension, because when arousal is high, attention tends to be directed outward for tasks such as scanning for new stimuli (Kahneman, 1973) - novel stimuli lead sensus to shift towards consciousness, and attention is focused on those stimuli. As the subject habituates, sensus shifts towards unconsciousness, and attention is freed to attend to internal processes or other external stimuli.

In this model, the position of the subject's mental state relative to the three dimensions determines the character of their experience (Waterworth & Waterworth, 2001). This state is dynamic, with shifts occurring due to a number of factors, which are not explicitly defined. However, the strong relationship between the focus and sensus dimensions, for instance, suggests how a sudden change in external stimuli could shift a subject from an unconscious state of sensus, to a conscious state while simultaneously shifting locus. Also, because the locus dimension is defined in terms of abstract or perceptual processing, it is possible to imagine that a task which demands a high degree of abstract thought would lead a subject to being absent.

The FLS model has been integrated into a complete psychological model, based on the notion that presence is strongly associated with consciousness (a conclusion also made by others such as Slater, 2002). In this Layers of Presence (LOP) model, presence functions on three separate but interactive levels of consciousness. At each level, presence is an evolved solution to some problem faced by the species during its evolutionary history (Riva *et al.*, 2004). Each of the three levels of presence acts to regulate the organism or to initiate action in the world (Riva *et al.*, 2004). The most fundamental problem which presence solves for an organism is distinguishing whether stimuli arise from inside itself, or from the environment (Waterworth & Waterworth, 2001); in humans, this sense has evolved to be significantly complex due the co-evolution of the mind, symbols and cultural artifacts (Riva & Waterworth, 2003). This has allowed presence in mediated environments.

The LOP model is defined largely in terms of neural activation patterns (no doubt derived from Damasio's habit of speaking of consciousness in the same terms). However, these neural patterns are so generally defined that it is possible to use this model as a set of psychological abstractions without specific reference to the brain. The LOP model derives from Damasio's (1999) concept of the self as having three layers, to argue that presence has three layers, each one corresponding to a layer of the self (see Figure 3.5):

1. *Proto self / proto-presence:* This unconscious part of the self contains the immediate state of the subject, including the current state of the sensory organs, as well as the internal state of the individual. Proto-presence represents the degree to which the subject can connect with the world at the most basic level – simple perception-action coupling (Riva & Waterworth, 2003). To be proto-present is thus to be effectively engaging with the world. This allows the subject to differentiate between the self and the outside environment.

2. *Core self / core-presence:* The core self is a conscious construct which is continuously updated with by both sensory information and past experience. It contains the current understanding of the subject's situation. The core self is highly dynamic, being constantly updated by changes in the external world (effectively implementing shifts in attention this way) and internal states such as mood and emotions (Riva & Waterworth, 2003). Core-presence is the outcome of focusing attention on a select subset of stimuli, to create a coherent mental picture of the current situation.

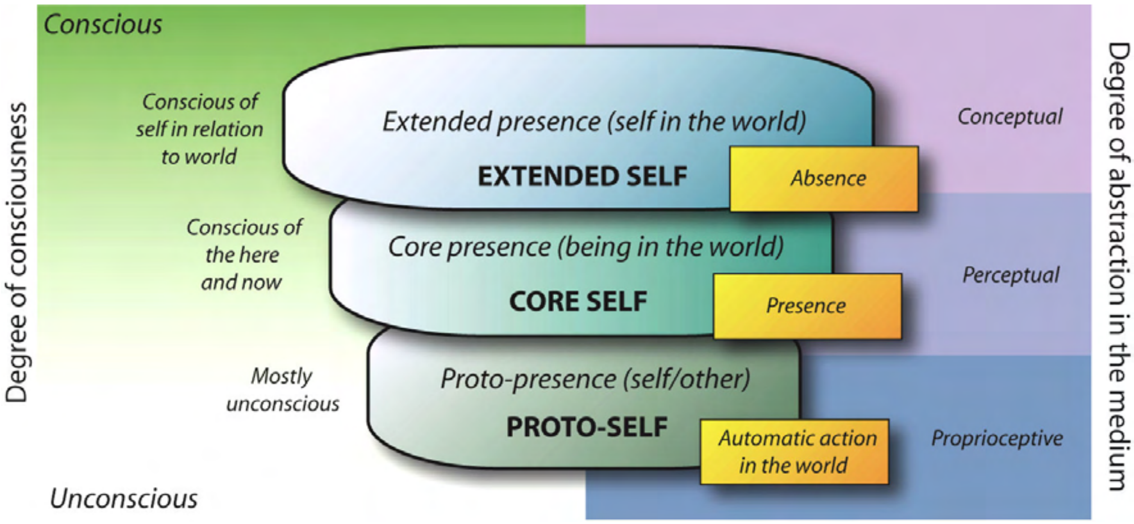


Figure 3.5: The levels of presence (LOP) model. Each of Damasio's layers of consciousness relates to a layer of presence (center ovals), the level of abstraction each responds to, and the type of behaviour which results from engaging each (yellow boxes). Maximum presence arises when all three layers are simultaneously engaged with the world.

3. *Extended self / extended presence:* This contains the most abstract processes, including invariants about the individual (such as biographical memory and personality). The extended self allows the individual to project their current state into the future, effectively making predictions and attributions not just about the individual, but about the environment also (Riva & Waterworth, 2003). The extended self plans, sets goals and creates expectations. Extended presence comes about by comparing the internal state of the extended self

(goals, predictions, etc.) with the environment's state to draw meaning from the individual's actions in the world. The feedback loop of extended presence is associated with achieving goals and extracting meaning about the environment.

3.3.3.2 Presence in the model

The three layers of presence, as with Damasio's three layers of the self, are not independent. They shift to respond to changes in the subject's internal state and to the external state of the world, and under particular conditions can achieve a high degree of integration (Riva & Waterworth, 2003). This integration (which is termed 'focused presence') is understood by the LOP model in terms of the focus, locus and sensus dimensions of the older FLS model. In the LOP model, the focus dimension (which determines whether the subject is focused on the environment or on the self) represents the degree to which the three layers of presence are aligned towards experiencing an external situation. When all three layers are integrated in this way (particularly when core presence is highly integrated with extended presence), the result is high presence; when they are not integrated, the result is absence. Presence can respond momentarily to a change in the environment (as in a break in presence) because proto-presence is highly sensitive to internal/external changes, and exists in the immediate moment. A small change in the environment or emotions can therefore trigger a change which will reduce the integration between the three layers and thus reduce presence (Riva & Waterworth, 2003). The locus dimension represents where the subject is situated experientially – in the real environment or in a mediated environment. Media provide a high degree of extended presence, as they are content-rich and abstract (E. L. Waterworth & J. A. Waterworth, 2003). Media, however, do not allow for direct perception-action coupling (they require an interface), and so proto-presence will not be engaged. The lack of integration between these levels means mediated environments will produce reduced presence when compared to real environments. Finally, the sensus dimension is related to the degree of arousal (Waterworth & Waterworth, 2001) which is passed from the proto-self (internal to the individual), through to the core and extended selves, allowing integration of the three layers and therefore high presence to occur more easily (Riva & Waterworth, 2003). Experiences which have a high degree of personal or emotional significance will begin by arousing the extended self (which understands the world at the most abstract level), and transmit downward through the core and to the proto-self, again facilitating the integration of the three layers of presence (Riva & Waterworth, 2003). The LOP model is thus capable of predicting some content effects in presence.

3.3.3.3 Summary of empirical evidence

As these are recent models, not much empirical evidence exists. The available data nonetheless seem to provide general support for the models. Unlike the three-pole model (Biocca, 2003), one cannot simply take evidence for the environment selection model as evidence for FLS/LOP, as there is no simple mapping between these models. Nonetheless, Riva *et al.* (2004) suggest how the existing corpus of findings about the immersion-presence relationship can be explained by the LOP model. At the top-most level, technology plays a very small role, as extended presence is largely internal to the subject. It involves drawing meaning and having one's predictions supported by the environment (Riva *et al.*, 2004). However, as one moves downwards towards the core layer, immersion has a larger role to play. The perceptually driven core self requires faster updates from the environment than the slower, conceptually heavy extended self. The VE must provide smooth, frequent updates to support core presence and allow its integration with the other two layers. Evidence for this comes from Meehan *et al.* (2002; , 2003) and Barfield and Hendrix (1995), who find that presence is related to update rate. Finally, extremely concrete proto-self canot operate with inputs which require decoding (Riva *et al.*, 2004). Realistic, high fidelity images are therefore preferable to iconic ones. This is supported by the bulk of evidence in support of the two-pole model (see 3.3.1.3 above).

The proto-self is also highly proprioceptive, so multi-sensory inputs which support proprioception will lead to the highest levels of proto-presence (Riva *et al.*, 2004). The evidence for this claim is also fairly clear – systems which use head tracking seem to produce higher presence (Bystrom & Barfield, 1999; Hendrix & Barfield, 1996a), as do systems which include haptic feedback (Meehan *et al.*, 2002; Sallnäs, 1999), and those which have interfaces that involve real reaching or other body movements (Schubert *et al.*, 2002; Slater *et al.*, 1995c). There is also evidence that physiological arousal (the sensus dimension of the proto-self) affects presence – Meehan *et al.* (2002) and Dillon *et al.* (2001) found that at least with for stressful and exciting VEs, change in heart rate correlates well with presence. Finally, in terms of extended presence, Riva *et al.*'s prediction is that non-immersive media will elicit lower presence than immersive media. Nunez and Blake (2003c) found that text-based VEs produce consistently lower SUS and PQ scores than desktop-based VEs; similar findings were reported by Lombard *et al.*, who found a difference in the expected direction between viewers of IMAX cinema, and viewers of small screen, black and white television (Lombard *et al.*, 2000a). Almost all these findings were published while the FLS and LOP models were under development, which makes this mostly post-hoc evidence, and does not suggest that FLS/LOP is any more powerful than the two-pole or environment selection models. Nonetheless, the theory's capacity to hold such diverse findings under a coherent framework is impressive. Two substantial tests of the FLS model have been published by Waterworth and Waterworth with colleagues (2003a; 2002).

The first of these (Waterworth *et al.*, 2002) describes observations of an interactive theater production evaluated in terms of the FLS model. The production (*Incarnation of a Divine Being*) was staged in a shared virtual space, with a chorus and chorus leader at one location, and the other actors (who are in fact also audience members) at different locations, such that most of the contact between participants was virtual. The piece was not scripted, but driven by the chorus and chorus leader who acted not only to drive the play, but also to elicit and mange the action and interactivity of the piece (Waterworth *et al.*, 2002). The VE allowed the participants (actors) to interact in the space by means of body-tracking in a stereopsis enabled large display (Waterworth *et al.*, 2002). The results of the experience were mixed, but give insight into FLS model. The chorus leader (a confederate of the researchers) had the greatest impact on subjects' experiences; many expressed surprise at the degree of their involvement. Most subjects begun the experience feeling anxious about taking part in such a public exercise, but this was replaced by a loss of self-consciousness as they began to interact in the experience (Waterworth *et al.*, 2002). In terms of the FLS model, the chorus leader increases the degree of focus, as his interactions demand the attention of the players, preventing outside stimuli from interfering in the experience. It also seems that participants experience a high degree of locus - they seem to have become part of the virtual performing group and situation, as evidenced by their loss of self-consciousness. This loss of self-consciousness also indicates a high degree of sensus, as subjects lost awareness of their own internal mental states. Although the importance of the chorus leader to the experience can be explained by other presence models, the loss of self-consciousness is more difficult for other models to address. The three-pole model has a strong emphasis on perceptual processing systems, which makes it difficult to explain changes in mental states; similarly, the two-pole model cannot explain the finding, except by reference to an 'acting as-if' explanation (Slater, 2003b), which essentially homunculizes the problem away rather than addressing it satisfactorily.

The second study which examines the FLS model directly (and the LOP indirectly) examined changes in presence and the estimation of time during two virtual experiences – one a field study, the other a laboratory study (J. A. Waterworth & E. L. Waterworth, 2003a). In keeping with the artistic sensibility generally found in the work of Waterworth and Waterworth, the VE used was unusually creative and designed to elicit a novel experience rather than to allow some particular task or practical purpose. The *interactive tent* is a low plexiglass half-tube, similar to a small tent, in which subjects lie and view back-projected images on the tent's surface. The tent also has a stereo sound system, with speakers on each side of the subject's head, and a subwoofer unit (J. A. Waterworth & E. L. Waterworth, 2003a). The subject can interact with the tent by shifting position and posture.

In the field study, the tent was used in an interactive art installation, *the Illusion of Being*. Subjects could control the form of the experience by moving their heads; left-right movements changed the experience from real to virtual (by shifting images and sounds from a realistic, filmed stream to an artificially generated stream) and up-down movements changed the experience from abstract to concrete (by shifting from images and sounds to written text and spoken words describing the scenes). The subject could therefore interactively select between four experiences (real/concrete, real/abstract, virtual/concrete and virtual/abstract). The content of the experience was constant (J. A. Waterworth & E. L. Waterworth, 2003a). Members of the public experienced the tent with no instructions or information given. After a seven minute experience, each subject was interviewed about their experience (J. A. Waterworth & E. L. Waterworth, 2003a). Most subjects did not realize the display changes were triggered by their head movements (many thought it was by means of measuring brain activity).

Most subjects reported changes in psychological state in response to changes in the form of the display, although with significant variation between subjects (J. A. Waterworth & E. L. Waterworth, 2003a). In general, subjects had a stronger sense of space during the concrete streams, and were more confused by the virtual streams. When asked how long the experience had lasted, almost all subjects underestimated the duration of their experience. Waterworth & Waterworth (2003a) conclude that the manipulation of abstractness and locus of the media form affects the character of the experience, as predicted by the FLS model.

To test these notions arising from the field study, *Illusion of Being* was adjusted for use in a laboratory study, which aimed to evaluate the effects of the experience on subjects' perceived duration of the experience. The subjective duration of an experience is related to how much mental work is done during that time: periods of high workload are experienced as longer, and periods of low workload are experienced as shorter (Waterworth & Waterworth, 2001; J. A. Waterworth & E. L. Waterworth, 2003a). The FLS model can explain this phenomenon. Experiences based on concrete stimuli (such as film) require less processing to decode, and because time estimation itself requires mental work, subjects to perceive them as taking longer. Conversely, experiences based on abstract stimuli (such as speech) will require more work to process, and will therefore be experienced as shorter. Because concrete experiences capture more focus, lead to a locus outside the body, and are more likely to stimulate sensus, they are more likely to lead to a focused presence experience (Waterworth & Waterworth, 2001). Concreteness is therefore predicted to correlate with both focused presence and length of time estimated. Time estimation can thus be used as an estimator of focused presence (J. A. Waterworth & E. L. Waterworth, 2003a).

The study used the same interactive tent, with the same four display streams, although subjects did not have control over which stream they experienced. Sixteen subjects experienced all four display stream, and were instructed to focus on the display rather than on estimating time. After each clips, subjects estimated the duration of the clip (J. A. Waterworth & E. L. Waterworth, 2003a). The subject's sense of presence during each clip was measured using eight items from the IPQ. Repetition effects were minimized by using a Latin squares design (J. A. Waterworth & E. L. Waterworth, 2003a). The IPQ scores largely matched the FLS model predictions. The concrete stream lead to more presence than the abstract stream, and the realistic stimuli lead to more presence than the artificial images. These findings are of course also predicted by the two-pole and three-pole models (as concrete and real images are essentially higher fidelity stimuli). The time estimation data were not as clear. Contrary to FLS predictions, there was no effect of media stream on time estimation. Regression analyses predicting time estimate from presence scores did show an effect, but only for one out of the four streams (the virtual/abstract stream). Waterworth & Waterworth (2003a) argue that the general direction of correlation in the other three streams suggest that the effect is general, although small. It may be that the small sample used reduced power and prevented the discovery of this small effect. As they stand the data do not support the notion that there is a relationship between estimates of experience duration and presence. Given that the one significant finding occurred in the case where some mental work was required to decode the content (the virtual/abstract stream), it could be that the relationship between mental workload and estimation is not linear, such that realistic environments require only a trivial amount of work to decode; or more likely, the decoding of realistic environments is handed off to specialized cognitive modules (Fodor, 1983), so that in effect ample mental effort is available for time estimation (Baddeley, 1986). Further work is required to resolve this issue, but as the FLS model stands, it cannot predict this lack of effect. The most convincing finding from this study in terms of the FLS model is the large degree of variability in presence scores, especially given that the tent experience was constant across all subjects. This variability may suggest that individual factors play an important role in the experience. However, given the small sample, it is possible that the variability is simple measurement error or other design artifact, and with a larger replication the effect may disappear.

3.3.3.4 Critical discussion of the model

An innovation of the FLS model is the inclusion of the sensus dimension, which provides an explicit role for physiological arousal in presence. This is important in the light of studies such as that of Meehan *et al.* (2002), which show that at in highly arousing VEs, presence varies with arousal. Also, the sensus dimension is useful in modeling the changes in arousal originating from shifts in attention, or from the arrival of a new stimulus into the perceptual field (Waterworth & Waterworth, 2001).

It is interesting that when FLS was developed into LOP, the sensus dimension was not explicitly converted into an emotion dimension, given that one of the major forces driving core presence is mood and emotion (Riva & Waterworth, 2003). Emotion cannot easily be modeled using arousal (high arousal could indicate anxiety or happiness, for example). It would be a logical step to explicitly include emotion in the LOP model rather than as a secondary force behind arousal. As the models currently stand, it is difficult to understand the exact contribution of sensus and arousal to the LOP model.

One way to understand the FLS model (and by implication much of the LOP model) is as an extension of the three-pole model. As Waterworth and Waterworth (2003b) point out, the locus dimension maps onto Biocca's physical/virtual axis, while the focus dimension maps onto the internal imagery/external stimuli axis. Waterworth and Waterworth (2003b) argue that the FLS model is more perceptual than the three-pole model, because the mental imagery pole requires conceptual processing. This argument is not convincing, because the processing involved in the mental imagery pole still involves manipulation of cognitive maps and perceptual representations of objects rather than abstract concepts. This becomes clear if one examines the underlying neural activation in direct perception as opposed to visual imagination. In functional MRI imaging studies comparing perception tasks to mental imagery tasks, several significant areas of the visual cortex activate in both tasks, and more importantly, similar shifts in activation occur when subjects change their mental images and when stimulus images are changed (Ganis *et al.*, 2004; Kosslyn & Thompson, 2003; Tong, 2002). Also, simultaneously giving a subject mental imagery and a perception tasks often leads to interference between the tasks, indicating that perception and mental imagery are functionally highly similar (Craver-Lemley & Reeves, 1992). This suggests that the three-pole model overwhelmingly emphasizes perceptual processing, while the FLS model, with its inclusion of the subject's body state in the sensus dimension takes a broader view. It is therefore expected that the FLS model would have more explanatory power than the three-pole model. This is supported somewhat by the results of Waterworth & Waterworth (2003a) discussed in 3.3.3.3 above, but the small sample size of that study limits its evidentiary weight.

A final criticism of this model is in terms of its measurement. Presence measurement is well-known to be a difficult problem (Nunez & Blake, 2003d; Singer & Witmer, 1999; Slater, 1999), and Waterworth and Waterworth recognize that this by taking a strong position that presence should be measured by objective means rather than by self report (J. A. Waterworth & E. L. Waterworth, 2003b). They suggest two methods of measuring presence (brain imaging and the time-estimate technique discussed in 3.3.3.3 above), which may indeed turn out to be valid and reliable measures. However, these suggestions raise serious complications for the FLS and LOP models, as each of these measures confound all dimensions and layers of presence into a single estimate. Given that the elements of the FLS and LOP models are internal to the psychology of the subject, they are quite difficult to manipulate. Using only one overall measure of presence (such as time estimation) it becomes quite difficult to validate the relative importance of each dimension or layer, and even harder to tease out the interactions between them. An ideal situation would include a measure of each of the three dimensions of the FLS model (or each layer of the LOP model), such that these can be isolated in studies. However, Waterworth and Waterworth do not offer a suggestion as to how this might be done; the specific validation of their models thus remains a fairly tricky proposition.

3.3.3.5 How the model explains the five problems

The book problem

Waterworth & Waterworth (2003b) explicitly state that the book problem is in fact incorrectly specified; it is not actually a problem. They argue that reading does not engage the senses, so the experience is not presence, but *"almost as if"* presence. They argue that the experience of reading is primarily engagement. This distinction recalls the four factor structure of the ITC-SOPI, TPI and PQ questionnaires (Lessiter *et al.*, 2001; Lombard & Ditton, 2004; Witmer *et al.*, 2005), which also separate *spatial presence* from *engagement*. However, as Lessiter *et al.* point out, these two factors are in most studies highly correlated (Lessiter *et al.*, 2001). This would seem to imply that the distinction, although theoretically quite clear, may not be so clean when one examines the data. Nonetheless, the FLS and LOP models are in fact capable of explaining the book problem. A book can engage the extended self quite effectively (as the reader can relate the text to their own experiences and predict how the story will develop), as well as the core self to some degree (the internal mental model created by reading the book will give the reader an idea of the present moment in the book in terms of spatial layout and states of the characters). However, a book will not engage the proto-self very effectively, as the story world exists only as mental representations which are internal to the reader. The outcome of this experience will therefore not be a particularly focused presence experience. The FLS and LOP models therefore are capable of explaining how books can lead to presence, and why they generally lead to less presence than immersive media, effectively solving the book problem.

The physical reality problem

The FLS model can explain the physical reality problem in terms of focus. The 'absence' extreme of the focus dimension describes the physical reality problem exactly. A subject who is close to absence on this dimension will not reach focused presence, regardless of their position on the other two dimensions. This carries through into the LOP model in the proto-self. A subject who is focused on internal processes will not achieve presence, regardless of the state of the other two layers of the self, as they will not be receiving input from the external environment.

The dream state problem

The FLS and LOP models are well able to deal with this phenomenon. In a dream, the locus and sensus dimensions are highly engaged (the dreamer experiences the imagined world directly, in a highly perceptual experience, and dreams are often extremely physiologically arousing). The focus dimension is also partly engaged, as dreams, although quite strange, are largely concrete. From the LOP perspective, all three layers of the self are engaged, with perhaps only the proto-self being under stimulated due to the disconnection of the motor system (E. L. Waterworth & J. A. Waterworth, 2003). A convergence of the three dimensions and three layers of the self is therefore possible, which predicts a high sense of presence for dreams.

The virtual stimuli problem

The locus dimension of the FLS model allows it to overcome the virtual stimuli problem in an elegant way. Rather than make a distinction between real and virtual environments (as the two-pole and three-pole models do), the FLS model makes the distinction between embodied environments and environments where action is mediated by hermeneutic relations, removing the need for more than one sensory information source outside the subject. The solution poses a few problems, however. Although the locus dimension explains embodied interactions well, it is not clear how it explains mediated interactions. For example, if a subject is sitting in front of two televisions (both mediated experiences), the subject can be more present in one than the other, and one can selectively switch between them (or be jerked from one to another by breaks in presence). The locus dimension does not explain such situations because it only recognizes switches between mediated and non-mediated spaces (as indeed do the two-pole and three-pole models). To completely overcome this problem, the model would need to explain how it is that from a single set of external stimuli, one can extract information to be present in many different places.

The inverse presence problem

The FLS and LOP models are able to explain this problem more fully than the two-pole and three-pole models, although they still do not provide a satisfactory explanation to this difficult problem. To explain inverse presence from a FLS or LOP perspective, one might begin by arguing that inverse presence occurs due to a lack of focused presence during a real-world situation; this may explain why subjects who experience inverse presence do not feel as if they are going through a normal presence experience. Working backwards, we would then assume that at least one of the three dimensions or layers of presence was not aligned with the others. Given that inverse presence is often associated with high anxiety or other physiological arousal (Timmins & Lombard, 2005), one would assume that the proto-self would be bombarded by demands to attend to the internal state of the organism, and similarly the focus dimension would be at the internal extreme. This might explain, from the FLS/LOPS perspective, high arousal situations lead to low levels of focused presence (even in a real environment). This fails to explain one essential aspect of the inverse presence, however – that the situation is experienced as mediated, to the degree that often the best explanation given by subjects is that it is like a movie or a television news report (Timmins & Lombard, 2005), and not like any other of a myriad forms of non-presence (such as daydreaming or plain absence). To explain this key aspect, one would require an associative link between the content of such experiences and the sense of mediation associated with such events. For example, one might expand Timmins and Lombard's (2005) memory based argument for the phenomenon to hypothesize that having experienced such events only through television, the subjective experience of mediation becomes associated with that type of content. When that content is encountered in the real world, the content becomes a retrieval cue for the subjective experience of mediation. However, no such mechanisms exist in the FLS or LOP models, essentially because the only type of learning or adjustment which they model is evolutionary. Neither the subject's previous experiences nor their media consumption history has been included in the model, and so such experiences are difficult to model.

3.3.4 The Measures, Effects, Conditions (MEC) model (Wirth *et al.*, 2007)

3.3.4.1 Description of the model

This recent model is largely cognitive, but includes some media and personality factors also. Presence occurs due to two processes: the construction of a spatial situation model, and the subsequent acceptance of that model as a viable hypothesis for interaction. (Wirth *et al.*, 2007).

The spatial situation model (SSM)

The SSM is a necessary condition for presence, and is therefore central to the MEC model. An SSM is described by Wirth *et al.* (2007) as a mental model of the space, with the following general properties:

1. *Completeness* – much a like a memory schema (Rumelhart & Ortony, 1977), the SSM is always complete, even at the earliest stages of exploring a new space (Wirth *et al.*, 2007), and develops more detail with exploration. Regardless of its detail, the SSM can be queried with a reasonable result ("reasonable" here is defined in terms of previous experience). On entering a new VE, a user's SSM of that apartment would be expectation heavy, based more on expectation than perceptual data. This allows even fragmented, incomplete sensory information to lead to a complete SSM. It is not stated whether semantic knowledge can make a contribution to the construction of an SSM equal to that of direct previous experience.

2. *Experiential coherence* – SSMs are derived from a combination of experiential knowledge and sensory information. They are therefore always coherent in terms of these two information sources (Wirth *et al.*, 2007). If sensory information arrives which significantly contradicts some experiential datum, the SSM is restructured to reduce the incoherence. This restructuring may be a slight change, or it may lead to the scrapping of the SSM to be replaced by an entirely new one.

3. *Idiosyncrasy* – As SSMs rely heavily on previous experience (Wirth *et al.*, 2007), it follows that they are highly personal to the subject. Although there will be commonality in SSMs constructed by different subjects (particularly for physical features such as size, colour, etc.), there will also be variation. Here a slight contradiction exists in the SSM concept – although they are defined as being idiosyncratic in this way, SSMs are also referred to as being more or less accurate than other SSMs (Wirth *et al.*, 2007). Given that SSMs are not an objective representation of a space, but a subject-centric representation of a space, it is not clear what accuracy of the SSM means, or what value there is in considering the objective mapping between the space and the SSM.

Phase I: Construction of the SSM

According to the MEC model, the first process in presence is the construction of an SSM. The first requirement for SSM construction is a subject's attention focused on the medium and its content (see Figure 3.6).

The MEC model is unusually sophisticated in its treatment of attention allocation, proposing two attention paths (Wirth *et al.*, 2007):

1. *Automatic attention allocation* – physical properties of the medium (loudness, brightness, etc) as well as unexpected or intense changes in the medium can lead to an orienting response, shifting attention to the medium (Posner, 1980). Because orienting responses occur quickly, they do not involve deep processing of the stimuli. Only physical features such as intensity and relative

position (and not content-related features) are likely to be processed (Posner & DiGirolamo, 1998). This is in line with the two-pole model which proposes that stimuli richness and multimodality are most important in eliciting presence (Slater *et al.*, 1994; Steuer, 1992). Of course, eliciting an orienting response alone cannot lead to sustained attention on the mediated environment. The stimuli must also support sustained attention by being at the very least comprehensible and moderately arousing (Posner & DiGirolamo, 1998). This agrees with Waterworth & Waterworth's (2001) notion of the importance of sensus (physical arousal) in presence.

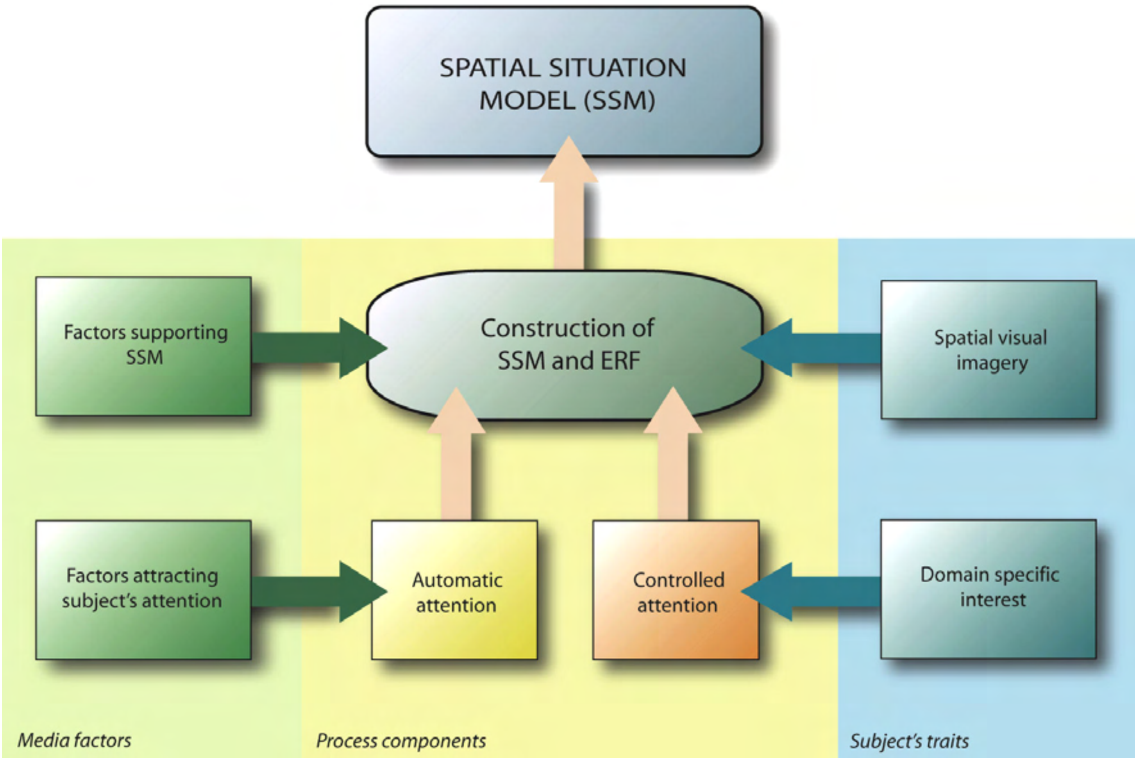


Figure 3.6: The first phase of the MEC model (SSM construction). Attention is allocated and a coherent model of the space is derived; both media factors (green boxes) and individual factors (blue boxes) affect this process.

2. *Controlled attention path* – to reflect the importance of motivation on ttention allocation (Bendiksby & Platt, 2006), the MEC model proposes several factors which explain the maintenance of attention on the medium. First of these is domain specific interest (DSI - Wirth *et al.*, 2007). DSI reflects subjects' increased motivation to attend to particular stimuli according

to their semantic content. Subjects who have their DSI engaged by a medium will find the medium interesting, and therefore willingly focus their attention on it. Apart from DSI, a number of other less significant factors affect the controlled allocation of attention, including fatigue, age, gender and emotional states (Wirth *et al.*, 2007).

These two attention allocation paths interact during processing. The form of the medium might elicit an orienting response, which would temporarily attract attention; then, based on the content of the medium, DSI and other factors might engage attention further. If during the experience a distracter competes for attention, form factors of the medium might allow for attention for remain focused. The degree to which each path contributes varies according to the medium being processed (Wirth *et al.*, 2007). For immersive media, the automatic path would be very active by virtue of the rich stimulus stream and lack of distracters; whereas for non-immersive media (such as books), the controlled attention stream would be most active, to compensate for the lack of stimuli which can produce orienting responses or ward off distracters.

Once attention is focused on the medium, an SSM will form automatically provided the medium represents a space. This process is moderated by two sets of factors (Wirth *et al.*, 2007):

1. *Spatial cues and media factors* – Spatial cues encoded in the medium are the most fundamental contributors to the construction of the SSM (Wirth *et al.*, 2007). This includes static cues (texture gradients, occlusion, spatial audio, etc.) and dynamic cues (motion parallax, stereopsis, Doppler shift, etc.). More cues lead to a more accurate SSM, although the term "accurate" in this context is not defined. These spatial cues must be presented in a coherent way (for example, with sound and visuals synchronized – Wirth *et al.*, 2007) in order for SSM construction to occur. Coherence is defined in terms of the subject's spatial knowledge of such environments (that is, the spatial cues should not obviously violate the subject's expectations for the environment). If one applies this discussion to media forms, it follows that strongly multi-modal, high fidelity systems will more easily allow for the construction of accurate

SSMs by the subject (a prediction which is largely in line with the two-pole and three-pole models).

2. *Spatial imagery and person related factors* –In order for an accurate SSM to be constructed from the available cues, the subject must have the ability to exploit the cues and cognitively process them (Wirth *et al.*, 2007). The best predictor of this ability is the subject's level of spatial visual imagery (SVI). High SVI subjects are better able to extrapolate cognitive structures from available perceptual data, even in the face of missing sensory information (Hegarty et al, 2002 in Wirth *et al.*, 2007). Interestingly, SVI has been linked to the ability to use metaphorical language (Tsur, 2002) and to the comprehension of poetical structure (Tsur & Benari, 2002). This may open an avenue for explaining the unexpected success of books at producing presence, and predicts a relationship between processing immersive and non-immersive media (Biocca, 2003).

Phase II: Selecting an ERF to be the PERF

Another important structure in the MEC model is the ego reference frame (ERF). ERFs are derived from SSMs, but encode a first-person perspective of the space defined by the SSM (Wirth *et al.*, 2007). ERFs are constantly updated as the subject moves through the space, and allows the subject to navigate or initiate action in the space (Franklin & Tversky, 1990). As ERFs are created during interactions with mediated spaces (Schneider *et al.*, 2004), Wirth *et al.* argue that it is therefore likely that subjects can maintain multiple ERFs (for example, one for the real world, and one for the mediated environment - Wirth *et al.*, 2007). Subjects will tend to switch to the ERF which is consonant with stream of stimuli which they are attending to, in order to reduce the resources required to process the environment (Wirth *et al.*, 2007). This stimulus-congruent ERF can be regarded as the primary ERF (PERF - Wirth *et al.*, 2007). When the SSM of the mediated environment is encoded as an ERF, and that ERF becomes primary, the subject experiences presence in the mediated environment (see Figure 3.7).

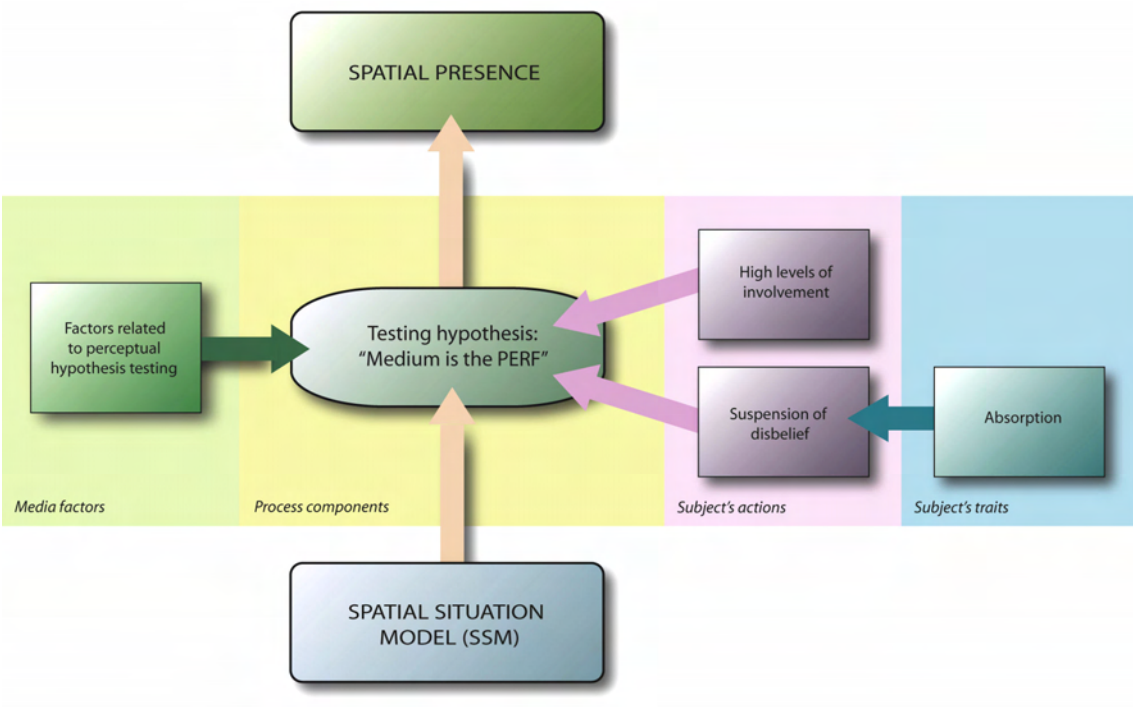


Figure 3.7: The second phase of the MEC model. Perceptual hypotheses are tested to select one ERF as primary, based on media factors and subject traits.

How one particular ERF becomes primary is explained by Wirth *et al.* by means of a hypothesis selection mechanism, based on the perceptual hypotheses theory of Lilli (Lilli, 1997; Lilli & Frey, 1993 in Wirth *et al.*, 2007). A perceiver always entertains multiple hypotheses about the scene, and perceptual information is used as evidence to confirm or disconfirm these hypotheses. The hypothesis with the most evidence is taken as true, and the perceiver behaves accordingly. As the perceptual information changes, new hypotheses may be formulated, and a different hypothesis may be selected as true.

Under this theory, it is easier to activate a hypothesis (prove it) than to deactivate it (disprove it). Hypotheses can be activated top-down (that is, by semantic priming and expectation) as well as bottom up (by perceptual data), although it is not clear what the relative contribution of each of these processes is. In the MEC model, presence defined as the state when the hypothesis "the mediated environment ERF is the PERF" is true (Wirth *et al.*, 2007). When enough evidence supports this hypothesis, and there is not an exceeding amount of contradictory evidence (as might occur during a break in presence - Slater & Steed, 2000), the hypothesis will be taken as true, and presence will occur. During hypothesis selection, the SSM is taken as a source of supporting evidence. A well defined, detailed SSM will support the hypothesis effectively, while a weak SSM (one which is consists mostly of conceptual information and is constantly contradicted by perceptual information) will not support for the hypothesis. These notions seem to be a more explicitly developed form of the environment selection model presented in 3.3.1.1 above (Slater, 2002; Slater & Steed, 2000), but the lack of reference to Slater's work in Wirth *et al.*'s paper and vice-versa suggests that the ideas were developed in parallel.

The hypothesis testing process is not entirely automatic. Two individual factors moderate the process at an abstract level: *involvement* with the medium content, and the *suspension of disbelief* by the subject. The MEC model sees involvement as non-critical acceptance of information from the mediated environment, which is strongly associated with the subject's motivation (Wirth *et al.*, 2007). Involvement has cognitive consequences (e.g. directing attention towards the medium or elaborating the stimuli to give them meaning), affective consequences (e.g. changes in mood or attitude towards the content) and behavioral consequences (e.g. selecting a particular action). As involvement is largely dependent on the content of the medium, it would seem that there is a weak relationship between involvement and spatial presence (as was argued by Slater, 2003a). Wirth *et al.* agree with Slater that involvement is not necessary for presence, but it can, under very particular situations, facilitate spatial presence for two reasons. First, a highly involved subject (who is highly motivated to experience presence) can lead to a subject willingly activate the ERF and therefore increase the probability of experiencing presence (Wirth *et al.*, 2007); second, highly involved subjects will automatically allocate more resources to processing the medium, which will leave less resources for processing competing stimuli. Less competing evidence means a higher probability that the appropriate ERF hypothesis will be selected and presence will result (Wirth *et al.*, 2007).

Suspension of disbelief is the term used by Wirth *et al.* for conscious strategies used to elicit or improve presence experiences. Suspension of disbelief is independent of involvement (Wirth *et al.*, 2007), although it seems to follow that it would be more effective if used by a highly involved subject. Wirth *et al.* see three components to suspension of disbelief: disabling the processing of contradictory or distracting stimuli; actively suppressing those contradictory stimuli which enter consciousness; and re-interpreting those stimuli which could not be suppressed as evidence in favour of the appropriate PERF hypothesis (Wirth *et al.*, 2007).

3.3.4.2 Presence in the model

As with the other models described, presence in the MEC model is a particular model state which arises under particular circumstances: when an SSM of the mediated space has been formed, and the hypothesis that the ERF encoding that SSM is the primary ERF is selected by the subject (Wirth *et al.*, 2007). Presence is a potentially fragile state, which can be interrupted by competing stimuli (as the PERF hypothesis could be selected out – Wirth *et al.*, 2007). A novel contribution of this model is the precise way in which it defines presence, which is possible due to the fact that this model limits itself to explaining spatial presence. The model does not include notions of presence as naturalness, or as engagement with the content, as suggested by Lessiter *et al.* (2001) and Lombard *et al.* (2000). It should however be noted that the relationship between involvement and spatial presence is indeed thoroughly discussed by Wirth *et al.* (2007).

3.3.4.3 Summary of empirical evidence

Although the MEC model is recent (published late in 2007), a significant amount of evidence exists suggesting its validity. Some pre-existing published work can be taken as evidence for the model, as in many ways this model expands the environment selection model by suggesting that subjects have at least two conflicting sets of stimuli, from which they select one in which to become present; and that a VE's fidelity, multimodality and the capacity to attract and hold attention predict presence (Slater, 2002; Wirth *et al.*, 2007). The MEC model also emphasizes the importance of consistency across stimuli. Evidence for this comes from Vinayagamoorthy *et al.* (2004), who simultaneously manipulated two aspects of scene realism: fidelity of characters in the scene, and fidelity of textures in the environment. The lowest presence scores were found when high fidelity characters were placed in the low fidelity scene. This suggests, as predicted by the MEC model, that presence is moderated by the fit between scene elements.

A few recent key studies have been carried out to test hypotheses which are specific to the MEC model. The first was a small study (n=26) by Gysbers *et al.* (2004), looking at the effect of number of spatial cues embedded in a text description of a space on both vividness of the SSM and spatial presence. Subjects read one of three passages describing a space: the first one contained a few spatial cues, the second contained many spatial cues, and the third contained many spatial cues plus instructions to imagine the space. The results showed, as predicted, that SSM vividness was related to number of spatial cues. The presence data, however, were inverted: more cues led to lower spatial presence. The authors interpreted this correctly to mean that SSM vividness is not a simple predictor of spatial presence. A close examination of the MEC model shows that SSM vividness is related only to the first phase. Presence will only occur if the SSM makes it into the PERF, which requires enough perceptual evidence support that hypothesis. It is likely that more spatial cues in the text *reduces* the evidence for that hypothesis, as more spatial cues in the text increases the probability of having perceptual contradictions (this is only plausible for text based environments, where the information is conceptual rather than perceptual). The study well demonstrates the complexity of the relationship between spatial cues, SSMs and presence (which one might argue justifies the complexity of the MEC model itself).

Evidence to support the distinction between the formation of an ERF from the SSM and the adoption of this ERF as PERF comes from the studies used to validate the MEC spatial presence questionnaire (MEC-SPQ) (Böcking *et al.*, 2004). 291 subjects under various media conditions completed an early version of the MEC-SPQ. A factor analysis revealed a reasonably strong three factor structure (explaining a little more than half of the total variance). The three factors were named by Böcking *et al.* as *self-location, possible actions,* and *cognitive involvement*. These three factors represent more the abstract factors in the MEC model: *self-location* is a measure of how an SSM becomes an ERF, *possible actions* measures the degree to which the mediated environment is taken on as the PERF (provided one accepts that any action is only possible if one positions oneself, hypothetically at least, in that environment), and *cognitive involvement* measures executive control over the adoption of the SSM as PERF. Although the use of factor analysis in presence theory has been criticized as a means of deriving theory (Waller & Bachmann, 2006), it should be noted that in this case the factor analysis was used not as an exploratory tool, but as a confirmatory technique (a preliminary version of the MEC model had already been published before this validation study; see Vorderer *et al.*, 2003).

An interesting feature of the MEC model is its fairly detailed account of the role of subject traits on presence. Noted that although the studies on the effect of personality variables on presence provide general support for the MEC model, they also support (although far less specifically), the LOP model. This is because the LOP model proposes that presence arises as a function of the layers of the self (Riva & Waterworth, 2003), which are presumably related to personality variables. Nonetheless, these studies provide stronger support for the MEC model than for the LOP model as the MEC model makes more specific predictions about these variables. One such study by Laarni *et al.* (2004) found that, as predicted by the MEC model, personality variables do have consistent effects on spatial presence. In particular *self-forgetfulness* had a noticeable effect. Self-forgetfulness is associated with easily losing consciousness of the self and of the passage of time when engaged in interesting activites (Kose, 2003). Interestingly, Kose (2003) uses the term “being in another world” (pp. 93) to describe high self-forgetfulness scorers, which immediately perks up a presence researcher’s ears. Similarly, the reference to the loss of awareness of time is reminiscent of Waterworth & Waterworth’s (2003a) study involving perception of time; it is therefore not surprising that this trait should predict spatial presence well.

A second study of interest is by Sacau *et al.* (2005), which examined three key personality traits posed by the MEC model – *domain specific interest, spatial visual imagery* and *absorption* (the first two are associated with the formation of the SSM, the third on is associated with suspension of disbelief and therefore with the adoption of the SSM as PERF). This large study (n=240 from four different countries) used four conditions: the first read linear text, the second read media-rich hypertext, the third watched a film, and the fourth navigated a three-dimensional VE. All four conditions encoded large, old buildings such as libraries or temples. Spatial presence was measured with the MEC-SPQ, which includes measures of spatial presence, domain specific interest, spatial visual imagery, and absorption (Vorderer *et al.*, 2004). The results showed domain specific interest and absorption are related to spatial presence as predicted (r=0.31 and r=0.19, respectively), but spatial visual imagery showed no relationship. When examining the effects of these three traits on the four types of media, the data showed that domain specific interest was a powerful predictor of spatial presence, only failing to predict spatial presence in the linear text condition. Absorption only predicted spatial presence in the media-rich hypertext condition. These comparisons may be somewhat blurred by the fact that the four media conditions were not randomized, but conflated with country (that is, all subjects in any one media condition came from the same country). Nonetheless, the finding is a confirmation for an important element of MEC, namely that interest in the content can enhance presence (presumably by means of allowing more control over attention - Wirth *et al.*, 2007), although a better test of this hypothesis would have been to manipulate an attention distracter during the study, as currently the path by which domain specific interest affects spatial presence is not clear. The fact absorption was a weaker predictor of spatial presence may be significant for the MEC model. While domain specific interest is thought to affect spatial presence at the stage of the formation the SSM, Absorption is thought to affect spatial presence at the stage of selecting the SSM as PERF. Recall that an accurate SSM is a necessary condition for its selection as PERF; this means that absorption affects spatial presence *after* domain specific interest has made its contribution. One would therefore expect a lower correlation between spatial presence and absorption, as it effectively functions as a moderator in the path between domain specific interest and spatial presence. This could be tested by explicitly running a path analysis with absorption as a moderator, or by having an explicit measure of suspension of disbelief (with which absorption should have a direct relationship).

3.3.4.4 Critical discussion of the model

The MEC model is well-defined with a sizeable amount of evidence supporting it. It is able to describe how perceptual data and conceptual data interact through a set of well-defined cognitive processes. Unlike the hypothesis-selection and three-pole models, the formation of the presence experience is not a spontaneous ‘black box’ phenomenon, but is posed as the outcome of two separate processes. These allow the model to explain failures to become present in the face of immersive media (either as a failure of the medium to attract attention at the first process, or as a failure for the SSM associated with the medium to become PERF), as well as breaks in presence (when a stimulus which is not consonant with the current SSM takes attention and either reduces the coherence of the SSM, or reduces the amount of evidence for it to be selected as the PERF). Finally, the MEC model is able to provide a lucid explanation of the interaction between medium content and spatial presence, by positing that attention allocation is moderated by domain specific interest. Even when compared to the more substantial FLS/LOP models, the MEC model is capable of generating very explicit hypotheses for empirical testing, and is able to link the higher levels of cognition (such as willing suspension of disbelief, controlled attention and domain specific interest) with very low level variables such as stimulus intensity and attention (the FLS/LOP models favour the higher level constructs, being more vague about lower level processes).

The MEC model also has some weaknesses which need to be mentioned. Firstly, the MEC model has been deliberately formulated as a model of spatial presence *in mediated environments* (Wirth *et al.*, 2007). This choice is particularly strange given the strongly psychological character of the model. From an evolutionary perspective, a model of presence must assume that whatever mechanisms lead to presence must have evolved long before the existence of mediated environments (Reeve & Nass, 1996; Biocca, 2003; K. M. Lee & Yung, 2005). A theory of presence should therefore be able to explain presence in real environments first of all. On careful reading of Wirth *et al.* (2007), it seems that the MEC model may indeed be able to adequately explain presence in real environments, if the importance of particular variables is modified. For instance, media related variables, such as the medium's ability to hold the subject's attention, would need to be reduced in importance.

A significant theoretical weakness in this model is related to the conception of presence in the model, and the use of the MEC-SPQ in studies validating it. The existence of the MEC-SPQ (which quantifies the subject's experience as a number based on a sum of responses to Likert type items) strongly suggests that the model treats presence as a continuously varying quantity. This notion is supported by how the spatial presence subscale scores are used in research (normally as covariates, or as outcomes to analysis of variance analyses – see for instance Böcking *et al.*, 2004; Sacau *et al.*, 2005). However, the model itself treats presence as a binary phenomenon, because presence is defined as the state when the SSM encoding the mediated world is *selected* as the PERF. If the scales are being used to support a binary concept of presence, one might expect a absolute score cut-off above which subjects are considered to be presence (similar to that done by Slater *et al.*, 1995c, where responses at the top end of his questionnaire as scored as 1, and all below as 0); however, the MEC-SPQ does not contain such a provision. This situation creates two fundamental problems: one, the definition and operationalization of presence are contradictory; and two, that a model which implements a binary concept of presence has been supported mostly by evidence derived from continuous measures of presence. This problem does not have a simple solution.

3.3.4.5 How the model explains the five problems

The book problem

Given that Biocca is included in the list of authors of the MEC model, one would expect the model to deal with the first three problems quite well, and this seems to be the case. In the MEC model, books and other non-immersive media can lead to presence as long as they are processed cognitively to produce an SSM (Wirth *et al.*, 2007). Recall that SSMs are created by a combination of sensory and conceptual data (hence their constant completeness). This implies that even with very little sensory input (as would occur in a book) an SSM can still occur. Also, there is no reason to think that a book cannot provide enough spatial information to allow for a detailed SSM, particularly if the subject has a high degree of spatial visual imagery. The difficulty lies in the book SSM being taken as the PERF. As books produce low levels of stimulation on a single modality, and they take considerable effort to decode, reading is easily interfered with by other stimuli. This makes it difficult for the 'book SSM as PERF' hypothesis to be maintained, and although presence can occur, it will likely be continually interrupted. Furthermore, the MEC model is capable of explaining why books generally produce presence for particular individuals better than for others. This occurs at both stages of the process – first, individuals with high spatial visual imagery will be able to construct a more accurate SSM, and second, individuals with high domain specific interest in the content of the book will be better able to control their attention though suspension of disbelief, and thus support the 'book SSM as PERF' hypothesis by eliminating support for rival hypotheses.

The physical reality problem

This problem is essentially one of mental effort being turned towards internal processing rather than to processing sensory stimuli. The MEC model includes a highly detailed description of the role of attention, so it copes with the physical reality problem reasonably well. If one assumes that the real world is encoded simply as another SSM and ERF (with extremely high levels of stimulus richness and multimodality, of course), then one can consider the amount of attention given to it and the amount of support for the 'real world SSM as PERF' hypothesis in the same was as a VE is considered. If the level of attention on the real world is low, very little support for the hypothesis will exist (due to a low degree of supporting evidence), and the hypothesis will be dropped. However, it is not clear what happens when the hypothesis is dropped, as the MEC model seems to assume that there will always be at least one other stimulus source which can lead to an SSM and ERF. This assumption exists because the MEC is a model of presence in *mediated* environments, and the real world is always assumed to exist as a replacement hypothesis. If the model is modified so that it is possible that no ERF is active, then this would explain the physical reality problem well. However, in its current form, the MEC model is not ideally suited to explaining this problem.

The dream state problem

This problem is similar to the physical reality problem, as it involves a situation in which internal processing is favoured while external stimuli are blocked out. This situation is possible under the MEC model, as discussed above, although it requires stretching the model somewhat. The interesting aspect of the dream state problem is the dream itself – it provides a source of high-level stimuli (not sensory, but perceptual) which can lead to presence. From the point of view of the MEC model, this is fairly straight forward. The dream provides a set of stimuli not unlike those produced by reading, and from these an SSM can be built if the correct cues are present. Unlike reading, individual factors such as spatial visual imagery will not play a large role, as during a dream the visual cortex is likely activated directly by the reticular activating system (Hobson *et al.*, 2000). If an SSM forms, then presence should occur if the perceptual hypothesis 'dream SSM is PERF' has sufficient evidence. Given that external stimulation is disconnected during dreaming (Waterworth & Waterworth, 2001), a great deal of supporting evidence can easily be collected for the hypothesis. However, the activation of the visual cortex during dreaming is largely random (Hobson *et al.*, 2000), so it is still possible for discordant images to appear and disprove the hypothesis.

The virtual stimuli problem

All of the models which have been discussed to this point (except perhaps the FLS/LOP models - Riva & Waterworth, 2003; Waterworth & Waterworth, 2001) have been unable to satisfactorily resolve the virtual stimuli problem, because they explicitly draw a distinction between a finite number of ‘streams’ or ‘worlds’ from which sensory data arises. The MEC model overcomes this problem completely by its use of SSMs. The SSM is an abstract structure which is composed of sensory and conceptual data which is selected by its content to be internally consistent from the sensory information available (Wirth *et al.*, 2007). There is no requirement that data arises from any particular source, or that the inclusion of some data, due to their origin, will reduce presence. Under the MEC model, every piece of spatial data (be it sensory or conceptual) is considered either as evidence in favour of or opposed to one or more SSMs and ERFs. If it fits (content wise) with the SSM or ERF, then it is evidence in favour of the perceptual hypothesis, and it increases the likelihood of presence; if it supports a different hypothesis, then it reduces the likelihood of presence. Note that this solution to the virtual stimuli problem is possible because presence is made an arbitrary state. One is said to be present if one particular ERF (of many possible ERFs) which we have designated as being of interest is supported as the PERF. In the MEC model, one is always present in some SSM (in that some SSM is always active in the current PERF); but “presence” is only counted when the SSM of the mediated environment we are investigating is the PERF.

The inverse presence problem

Although the MEC model is able to explain the other four problems reasonably well, it unfortunately does not deal with the inverse presence problem well. In the MEC model, presence is defined as simply the switching to a PERF previously defined as interesting. Mediation is considered only as a factor which reduces the quality of sensory data, and makes the subject less likely to attach and hold attention to those stimuli. This cannot capture the essentials of inverse presence which are the qualia of a mediated experience, and the triggering of particular expectations based on the content of the environment (Timmins & Lombard, 2005). This cannot be modeled unless some system of automatic, associative memory is included, such that particular features of a situation can trigger off contextually dependent expectations (e.g. fire and explosions must mean that Arnold Schwarzenegger is about to enter the scene, and I feel like eating popcorn). The MEC model does include aspects of memory and previous experience - these are the components of the library of spatial experiences which subjects use to interpret sensory stimuli as spaces. But these contain only spatial information, and are not linked to episodic memories of previous experiences in similar spaces which may lead to the expectations characteristic of inverse presence. One can defend the MEC model in this regard by stating that it is a model of *spatial presence*, and its lack of power in explaining inverse presence comes from an intentional design choice rather than an inherent model limitation. This may indeed be the case, but if it is so, the MEC model will also have problems explaining other phenomena related to the qualia associated with particular spaces (such as a feeling of awe on entering a cathedral).

3.4 Conclusion

This chapter has summarized four of the current significant families of presence models on a common framework to aid comparison (see the summary table in Appendix G). Most models have some roots in the two-pole model, and thus benefit from the large body of empirical support of that model. Generally speaking, the two-pole/environment selection and the MEC models have the most significant body of empirical evidence, and the FLS/LOP models have the least. In terms of the five problems, no one model was able to convincingly deal with all five. The two-pole/environment selection model is the weakest in this regard, dealing only with the dream-state problem. The strongest in this regard was the MEC model, which could deal (at least partly) with all except the inverse-presence problem. The following chapter will propose the CLICC model, which aims to distill the strengths of the four families of models discussed in order to deal with all five problems, while remaining consistent with previous empirical findings.

Chapter 4: The capacity limited, cognitive constructionist model of presence (CLCC)

This chapter describes the capacity-limited, cognitive constructionist model (CLCC) of virtual presence. The chapter begins by giving a broad justification for the architecture used (sections 4.1 and 4.2), and then moves to describing the components and structure of the model (section 4.3). Finally, it discusses how presence exists in the model (section 4.4), and illustrates the predictive power of the model by explaining, among other theoretical points, the presence-immersion relationship (section 4.5) and the five presence problems defined in 3.1 in chapter 3 (section 4.7).

The CLCC model contains a complex modular structure (see 4.1 below), and it models presence as a dynamic model state, which can be understood as consisting of two main effects which occur simultaneously and interactively:

1. The creation of a semantic (themed) bias which permeates the model
2. The construction of the current environmental situation in working memory

This interaction keeps the model relevant to the environmental situation, in as cognitively efficient a manner as possible (this is necessary due to the small capacity of human working memory – Baddeley, 1986). Once a particular construction or understanding of the environment has been arrived at, the model is present in that environment. The model will then try to maintain that construction by filtering out perceptual information and inhibiting the influence of top-down data which are semantically irrelevant. If it should occur that the current construction is no longer relevant to the environment (either due to the existence of a lot of contradictory data, or because of an unexpected change in the environment), a process of reconstruction will occur, which will make a new, meaningfully coherent construction of the environment. The name of the model describes this basic operation: it is *capacity limited* due to the cognitive constraints of human information processing, and *cognitive constructionist* due to the semantic construction of a meaningful, coherent mental structure of the environment by the subject.

4.1 Structural basis

As suggested by Biocca (2003) and Lee (2004), the mechanisms which give rise to the presence experience must serve an evolutionary purpose more fundamental than presence in virtual environments. These mechanisms exist to provide the subject with an up-to-date mental representation of the environment, into which are encoded learned potential courses of action. Because evolution tends to develop systems which are brutally efficient (Tooby & Cosmides, 1990), this is achieved within the constraint of highly limited capacity, such that all parts of the system are always using all available resources to process the scene and infer possible actions. Because the system has evolved to deal with stimuli which encode real environments, it is possible to fool it, provided stimuli are presented in a form which somehow stimulates the appropriate cognitive modules (Lee, 2004).

In order to the 'lean and mean' character of evolved cognition, the CLCC model of presence is compatible with current findings in cognitive science, while being able to predict and explain presence phenomena. The structure of the model largely follows the information processing models of cognitive psychology. Specifically, it follows the *stages of processing* model of Atkinson and Shiffrin (1968), and to some extent the *levels of processing* model of Craick & Lockhart (1972). Although there has been some controversy about these models of memory, a large number of empirical studies over the past thirty years (including several meta-analyses), suggests that these are robust and well supported models of semantic processing (Conway, 2002; Decoster & Claypool, 2004). In the tradition of the stages of processing model, the CLCC model proposes that the cognitive architecture which leads to presence is a set of discreet processing modules which take data as input from some nodes, transform that data, and output it to another set of nodes for further processing. Unlike classic stages of processing models, however, the CLCC model operates due to emergent properties of the structure, which arise due to the interaction of the data content, and how it flows around the entire model. Considering presence as the emergent property of a complex system is common in extant presence models - for instance, it has been used in the LOP model (Riva *et al.*, 2004) and MEC model (Wirth *et al.*, 2007) discussed chapter 3. The role of emergence in the CLCC model has also been inspired by the levels of processing model (Craik & Lockhart, 1972), which states that in a cognitive processing system, information is encoded and processed at multiple levels of abstraction, from perceptual to conceptual (Craik & Lockhart, 1972). Furthermore, the levels of processing model states that synchronous processing of data across several levels of abstraction leads to an enhancement of overall processing performance (Craik & Lockhart, 1972).

Many extant presence models are designed to only explain presence phenomena (such as the presence models of Biocca, 2003; Slater, 2002; Wirth *et al.*, 2007). The CLCC model however has been designed to closely reflect existing general purpose human information processing models, as is done by the LOP model (Riva *et al.*, 2004). This approach has two advantages; one theoretical, and one pragmatic. Theoretically, it satisfies Biocca's (2003) and Lee's (2004) evolutionary requirement that presence should not be considered as a unique response to mediated environments. More pragmatically, using a well-understood architecture as the basis for a model reduces the need to empirically validate the overall structure of the model, allowing empirical investigation to focus on specific aspects of the model which are pertinent to presence.

4.2 Cognitive constructionism

A central axiom of the the CLCC model is that each subject constructs their own unique experience of an environment by applying previous knowledge of similar situations and by interacting with the current situation (Bruner, 1990). In this way, the subject can operate in novel environments, but in an adaptive way which is constrained by previous experience. This idea is derived from educational (and particularly Marxist) theories (see for instance Luria, 1974; Vigotsky, 1978), but also exists in cognitive psychology as the idea that goal-directed behaviour can only occur when there is an interaction between bottom-up and top-down information. Bandura (1986) referred to this as *reciprocal determinism* – the subject, as an active agent, changes the environment, and is simultaneously motivated by those changes to respond in particular ways. Bandura's explanation operates at the level of the individual, but similar concepts exist at lower levels. For instance, Rumelhart and McClelland's layered, competition-based connectionist networks only reach a stable state (and thus complete computation) when activation has been resonated between the top-most and bottom-most levels of the network (Rumelhart *et al.*, 1986).

Inclusion of constructionist principles allows a presence model to simultaneously explain cognitive phenomena (such as the role of attention allocation or the importance of multimodality; see Sallnäs, 1999; Slater & Steed, 2000) and subject-environment interaction (such as the role of passive haptics or successful navigation; see Barfield & Weghorst, 1993; Meehan, 2001).

Three essential features of constructionism (Bruner, 1990) which may be useful to a presence model are:

1. *The subject is an active agent in the world* – a subject's normal mode of being in an environment involves exploration and interaction. This is explicit in some models of presence, such as the argument for embodiment present in Biocca's model of embodiment (Biocca, 1997), and the Locus dimension of the FLS model (Waterworth & Waterworth, 2001). Not all models of presence recognize the subject's agency in the environment however. Notably the earlier versions of the environment selection model (Slater, 2002), which see the subject as passively selecting and processing information about the environment, with the presence experience arising more or less automatically when the correct conditions are met (Biocca, 2003; Slater *et al.*, 1995c; Witmer & Singer, 1998). Recent developments of the environment selection model (Slater, 2002) have emphasized the subject as actively participating in their experience, suggesting that there is a fair degree of consensus about the importance of agency in presence.

2. *The subject interactively creates their experience* - under this principle, the subject's experience of the space depends on their actions and goals within that space (Bandura, 1986). This notion was first applied to telepresence by Sheridan (1992b) with his elegantly simple model of in-the-loop human-machine interaction and telepresence. Recently, it has been extensively adopted by the MEC model (Wirth *et al.*, 2007), which emphasizes the role of the SSM (an interactively constructed model of the environment). The importance of a subject's interactions in the VE on their presence experience has been an area of active inquiry. Schubert *et al.* (2002) showed that subjects' perceptions of their degree of interaction in the VE positively predicted

presence. Other studies which have examined the relationship between presence and task performance or task completion (which require interaction in the VE), have shown that, for instance, navigation (Barfield & Weghorst, 1993), self-rated ease of task (Bystrom & Barfield, 1996) and self-rated task performance (Romano *et al.*, 1998) all correlate with presence in the expected direction.

3. *The final product is meaningful to the subject* – as the experience is constructed by the subject by combining the current situation with previous experience, it is essentially idiosyncratic (Bruner, 1990). This implies that the content of a VE will interact with the subject's previous knowledge during the presence experience. Although Slater has argued against the role of content in the presence experience (Slater, 2003a), the MEC model (Wirth *et al.*, 2007) proposes that the subject's domain specific interest (which predicts their interest in and experience with the VE content) plays an indirect role in the presence experience. Similarly, the FLS model (Waterworth & Waterworth, 2001) proposes that how arousing the content of the VE is will partly determine the nature of the presence experience. There is very little empirical evidence of how the meaning given to a scene affects presence in that scene. An interesting study by Nowak *et al.* (2006) showed that for violent video games, presence was predicted not by the objective degree of violence in the game, but by the degree of violence *perceived* by players, suggesting that the meaning attributed to the scene may in fact contribute to the presence experience.

4.3 Structure and components of the model

The CLCC model has the same basic architecture as the stages of processing model (attention filter, short-term storage, long term storage – Atkinson & Shiffrin, 1968), with some important differences – see figure 4.1 below. First, unlike the stages of processing model, the data paths between components in the CLCC model are one-way – that was done simply to make more explicit the stages of data transformation. Second, the CLCC model is not only a model of memory encoding and retrieval. Although memory plays an important role in the CLCC model, it also includes aspects of perception and motor control. Finally, while the stages of processing model describes only the subject, following the arguments of Floridi (2005) and the theoretical suggestions of Biocca (1997) and Schubert *et al*. (2002) the CLCC model is epistemic - it includes aspects of the environment, and provides a mechanism for feedback via the subject's perception of their own actions in the environment.

The components of the model are as follows (described from bottom to top):

4.3.1 External stimuli via sensory cortices

Description: This is the entry of bottom-up data into the model. The data are in the form of unprocessed neural stimuli from the retina, the auditory nerves, etc. This node (which includes basic perception done in the sensory cortices) converts raw neural stimuli into objects. For example, visual stimuli are converted into a 2½ dimensional sketch (Watt, 1988), and motion is identified by optic flow on the retinal field (Cavanagh & Mather, 1989).

Inputs from: The physical environment (including all immersive systems, real events, the body itself, etc), which provides physical stimulation.

Outputs to: The stimulus attenuator for selection of relevant stimuli.

4.3.2 Stimulus attenuator (attention)

Description: The primary purpose of this node is to ensure that capacity limits of the system are not violated (Baddeley, 1986; Posner, 1980). This is done by selecting a subset of stimuli from those available at the sensory cortices. Stimuli are selected as a function of their relevance to the current semantic bias of the entire system (Lavigne & Denis, 2001; Maxfield, 1997). However, this selection is not a simple linear function. The probability of a stimulus being selected for further processing is highest at both extremes of congruency to the current system bias. Highly congruent stimuli, which support the current bias, are selected, and incongruent stimuli are excluded to maintain semantic coherence. However, in order to remain sensitive to significant changes in the environment, highly incongrious stimuli (either in terms of physical properties such as brightness, or in terms of semantic difference) are highly likely to be selected for further processing (Treisman, 1969), which would result in a violation of semantic coherence, and a reconstruction (see 4.3.11).

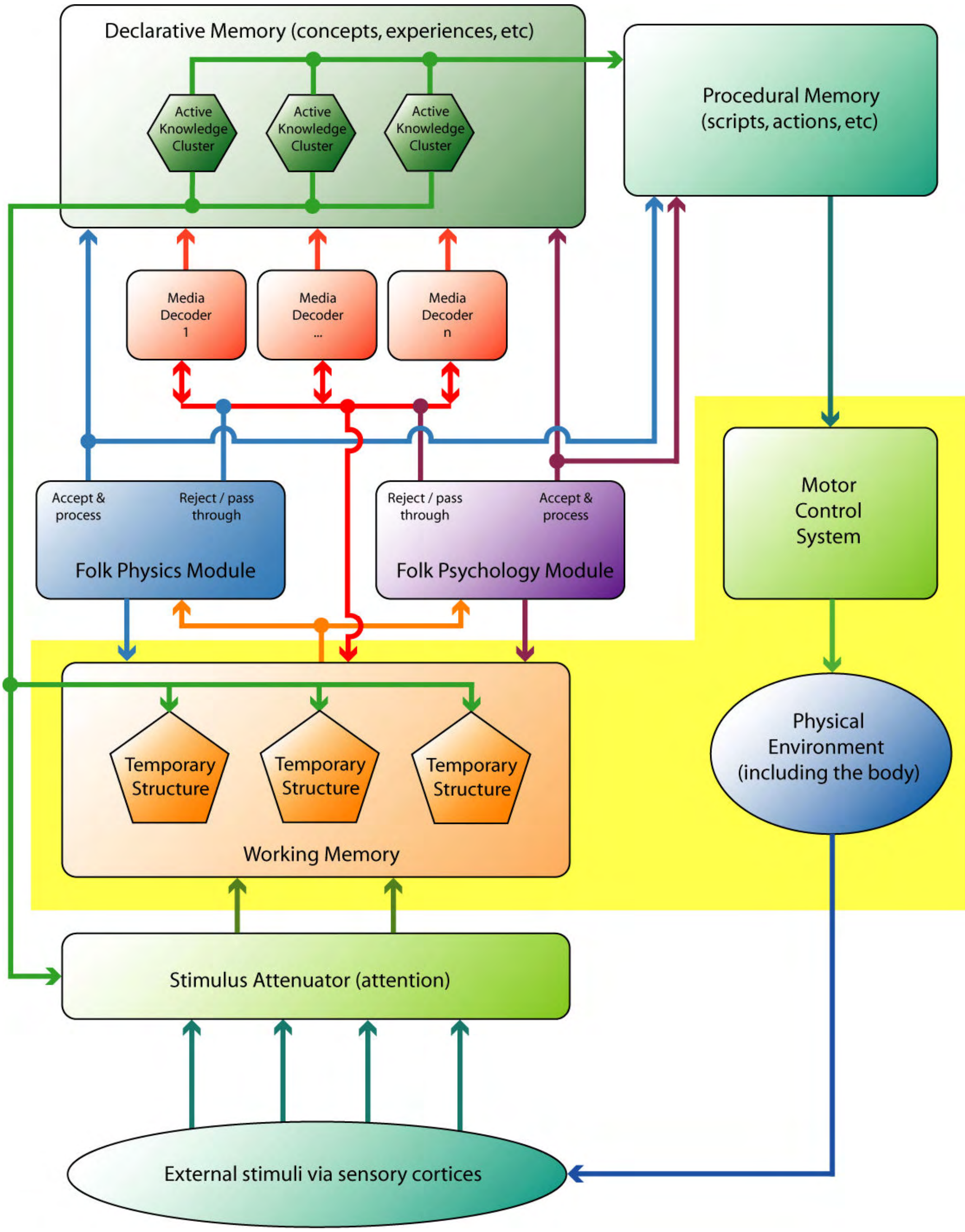


Figure 4.1: The structure of the CLCC model. Information flows only in the directions indicated by the arrows on the information paths (see section 4.3. for discussion). The three nodes in the yellow shaded region are the only nodes at which measurement is possible (all other nodes are cognitive abstractions implemented by unknown neural mechanisms).

Attention has been widely discussed in the presence literature (Riva & Waterworth, 2003; Slater & Steed, 2000; Waterworth & Waterworth, 2001; Wirth *et al.*, 2007; Witmer *et al.*, 2005). In the CLCC model, the stimulus attenuator models attention as an extremely simple mechanism which reduces the number of incoming stimuli while requiring only minimal processing to make the decision (in the tradition of Treisman, 1969). This mechanism allows the attenuator to model breaks in presence. Although the term 'break in presence' strictly applies to a subjective experience without reference to a particular cause, the most non-controversial uses of the term apply to situations where a subject suddenly experiences an expected stimulus, such as the tug of an HMD cable (Slater & Steed, 2000), a rendering glitch by the system (Brogni *et al.*, 2003) or an experimentally introduced artifact on the display (Vinayagamoorthy *et al.*, 2004). All of these situations pose that a stimulus which is incongruent to the semantic bias of the system (see 4.3.6 below), demands attention to itself and breaks the subject out of the experience. It is hardly surprising that a tug of a cable should lead to a break in presence; however, what is interesting about a break in presence is that given the current achievable levels of fidelity they do not *constantly* occur. In current VE systems, subjects must always deal with a large number of competing and conflicting stimuli. This is because the stimulus attenuator, with its automatic tendency to exclude moderately incongruent stimuli removes these artifacts from processing before they can adversely affect the experience. However, a sudden intense stimulus (especially if it is highly incongruent to the semantic bias of the system) will be selected for further processing, leading to a reconstruction and thus a break in presence. Note that not all intense stimuli will lead to a break in presence (this will only occur if that stimulus leads to a reconstruction) – if the intense stimulus is semantically related to the VE (a roaring jet passing overhead in a VE of an airport), then it may strongly reinforce the current construction.

Inputs from: The sensory cortices (from which it selects data), and from the active knowledge clusters in declarative memory which implement the semantic relevance bias.

Outputs to: Working memory, to provide perceptual (bottom-up) data for the construction of temporary structures.

4.3.3 Working memory

Description: This is a temporary buffer used during processing which maintains a mental representation of the current situation (Baddeley, 1986, , 1998). The CLCC model makes use of the working memory model of Baddeley rather than the more recent models such as that proposed by Cowan (2001) and Oberauer (2002) due to the extensive use of Baddeley's model in human factors research which has seen application in the presence field - for instance, see Bysrom et al's IPP model of performance in VEs (1999a), or Stanney *et al*'s review of human factors issues in presence research (Stanney *et al.*, 1998). The contents of working memory are accessible to consciousness (Baddeley, 1998; Rumelhart *et al.*, 1986), which makes it a potential site for presence measurement. Working memory's capacity is highly limited – only between five and seven meaningful chunks can be stored (Baddeley, 1986; Cowan, 2001). Working memory is not a unitary store – it has two known subsystems, which operate on specific modalities. The *phonological loop* is used for verbal (and auditory) information (Baddeley, 1986), and its primary purpose is processing of speech (Baddeley, 1998). It also plays an important role in reading (Carpenter & Just, 1989). The second subsystem is the *visuospatial sketchpad*, which is used to process visual and spatial data (Baddeley, 1986). It also plays an important role in the simulation of physical events such as predicting where thrown objects will fall (Cowan, 2001). Each of these systems has its own store of capacity, such that loading one will not affect tasks which make use of the other (Baddeley, 1986, , 1998). For example, a navigation task (which loads the visuospatial sketchpad) and remembering a set of numbers (which loads the phonological loop) will not interfere with each other. However, if one engages in remembering a set of numbers and reading a piece of text simultaneously, then performance in one or both of the tasks would be adversely affected, as the tasks are competing for verbal working memory capacity. Most cognitive processes require some working memory (with the exception of processes which have dedicated neural circuits, such as face recognition; see Andreasen et al., 1996).

Working memory can be usefully applied to presence, but has hardly received any attention in the literature. Baddeley's working memory model (1986; , 1998) has been used to explain phenomena which, from a micro-cognitive perspective, are similar to presence. These phenomena include reading comprehension (Carpenter & Just, 1989), and spatial navigation (Garden *et al.*, 2002). These tasks involve the processing of a subset of external stimuli so as to decode some meaning (be it spatial or otherwise) to allow further inferences about the space or about action in the space. If it is true that presence involves both perceptual and conceptual processing of an environment (Waterworth & Waterworth, 2001; Wirth *et al.*, 2007), and given that all processing requires some working memory (Baddeley, 1986), then it follows that for presence to occur, some amount of working memory will be required to process the environment.

How working memory is allocated to a task is a complex process. Simply having more information to process does not imply that more working memory will be allocated, or that more working memory will be needed. Data which are meaningfully related are automatically *chunked* into fewer, more abstract complexes, effectively freeing up working memory space (Baddeley, 1986). Thus, an immersive display which renders a large number of highly correlated variables (both in terms of perceptual synchronization and semantic organization) may require little working memory to process, as the information can be easily chunked. If a latency were to develop in the display of one of the channels (for instance, in sound rendering), then the temporal discrepancy would prevent the chunking of sound together with the other variables, and thus more working memory would be required to process the scene. This same mechanism can explain why stimuli which come from 'outside' the virtual environment (as with the radio position manipulation discussed in Slater *et al.*, 1995c) can reduce presence – due to the spatial discrepancy, these stimuli will not chunk with the stimuli 'inside' the virtual environment, and will thus require more working memory to process. If enough of these anomalies abound, then they will begin to impinge on the working memory which is necessary for successfully processing the VE. This reduction in the amount of processing focused on the VE will then lead to a reduction in the sense of presence.

Inputs from: The stimulus attenuator, which provides relevant perceptual (bottom-up) stimuli for processing. Secondary inputs are taken form the media decoders, and the folk psychology and folk physics modules, which use working memory as a scratchpad during processing.

Outputs to: To the folk psychology and folk physics modules, for semantic decoding.

4.3.4 Folk physics and folk psychology modules

Description: These modules recognize and respond to the contents of working memory. The folk physics module infers physical properties such as mass, velocity and spatial arrangement from these stimuli. It is also responsible for and accessing mental models and cognitive maps (Plotkin, 1998). The folk psychology module infers psychological properties, mental states and intentionality; it also responds to facial expressions, and is used in inferring emotion (Plotkin, 1998). As these modules are evolved for a particular purpose, they are highly specialized, and only respond to highly specific stimuli. The folk psychology module may allow for the bridging of individual forms of presence (the as presented in this form of the CLCC model) with social forms of presence.

The inclusion of these modules in the model is for two reasons: First, due to the empirical evidence from cognitive science that these modules are highly automatic and lead to important effects on many aspects of cognition including perception and memory (see Plotkin, 1998; Pinker, 2004); and second, on the suggestion on Lee (2004) who argued that these modules are involved in presence due to their tendency to automatically process even stimuli which crudely represent the objects they have evolved to process (this argument in detailed in section 2.2.2 in chapter 2).

Inputs from: Data are taken from the active temporary structures in working memory. These are tested against a minimal set of features to determine their suitability for processing in this module.

Outputs to: This module has two main output paths. The first path is used when the stimuli are not of the type processed by the module – in this case, the data is passed through to the media decoders for processing there. If the data are of the type to be processed, the data may be used to trigger a reflex action. For instance, the folk physics modules may trigger leaning into a curve during a virtual car ride (Freeman *et al.*, 2000), while the folk psychology module may trigger automatic following of eye gaze (Baron-Cohen, 1995). When an automatic response is required, the data are immediately passed to procedural memory for selection of a motor program for execution; otherwise, they are passed to declarative memory for semantic processing. A minor feedback path also exists from each module back to working memory, which allows the modules to make use of working memory as temporary storage during processing.

4.3.5 Media decoders

Description: Decoding the contents of working memory into semantic meaning is complex because the specific operations required vary depending on the medium used. Decoding a photograph of a room requires different cognitive processes to decoding a verbal description of that room, even though the final semantic products will be similar. In the CLCC model, decoding of media is done by learned cognitive modules (one per medium) whose inputs are data from working memory, and whose outputs are abstract representations of the content of the medium in declarative memory. Each of these media decoders is a collection of strategies and processes for decoding one particular medium. There are many possible decoders – a writing decoder (after evidence from Carpenter & Just, 1989), a film decoder (after evidence from Bordwell, 1989), a diagrammatic decoder (after evidence from Allmendinger, 1998), and so on. The decoder will only exist in a given subject if that person has learnt how to decode that medium. When a new set of perceptual stimuli are considered for processing, the appropriate media decoder is selected on the basis of a small set of key features (for instance, the basic shape of letters in declarative memory might trigger the writing decoder). Once the media decoder has been activated, it proceeds automatically with decoding the stimuli. If a media decoder attempts to decode the wrong type of medium (for instance, in the case of a picture being embedded in text as might occur in a magazine), then the error becomes a signal for the selection of a different decoder.

Media decoders, as processing units, require working memory. How much working memory a particular media decoder requires is a complex question. It seems reasonable to suggest that some decoders will require more working memory than others. For instance, images and video are relatively easy to decode (partly because there are dedicated neural circuits, in the visual cortex and other areas, specialized for this task - Krubitzer, 2005), whereas writing requires more resources to decode (a first a visual pass is required to decode individual letters and words, and a parallel second language decoding pass to decode the meaning of the sentences as a whole; Carpenter & Just, 1989). It also seems reasonable to suggest that some media decoders become more efficient with practice. For instance, reading requires time to learn and generally improves with practice, eventually becoming almost effortless (Carpenter & Just, 1989). Similarly, some film genres make use of conventions which must be learned at first, but are later decoded with little effort (Bordwell, 1989). From a working memory perspective, this increase in efficiency and associated sense of effortlessness come from a decrease in the amount of working memory used by the decoder as it becomes more efficient in chunking data (Baddeley, 1986).

Inputs from: Data are taken from the pass-through outputs of the folk psychology and folk physics modules.

Outputs to: The decoders have two outputs: One is to declarative memory, which is used to activate semantic meaning (the product of decoding). The second path is a feedback path to working memory used during processing. Media decoders can also shuffle data to other media decoders along a parallel bus. This will occur in two situations: when a decoder takes data but finds it cannot process it, or when decoding complex media (such as text embedded in a film), which requires two decoders to simultaneously process a scene.

4.3.6 Declarative memory

Description: This memory system stores two types of information: semantic information (including concepts and the relationships between concepts – Squire, 1999) and experiences (including the temporal organization and personal significance of experience – Squire *et al.*, 1993). Unlike working memory, declarative memory is effectively unlimited in terms of capacity (Squire, 1999). Declarative memory is used in the CLCC model in preference to the more specific subdivision into semantic and episodic memory suggested by Tulving (1995) for reasons of parsimony – there is currently not enough data to infer separate roles for semantic and episodic memory in presence.

When processing an environment, declarative memory associates semantic meaning to percepts, and contextualizes with respect to the subject’s previous experiences.

Various models of declarative memory encoding exist, but the preferred concept in the CLCC model is that of schemata (Rumelhart *et al.*, 1986). A schema is a concept encoding structure which is modified by experience to contain default information (Rumelhart & Ortony, 1977). This allows the inference of missing information. For instance, if in a restaurant I see a man wearing a black bowtie and holding a notepad, I can infer that this man will take my order and I can pay the bill to him at the end of the evening (by activation of the 'waiter' schema). Schemata are connected in semantic networks, such that activation of one schema automatically leads to the activation of related schemata, thus creating semantic contexts (Rumelhart *et al.*, 1986). These small networks of activated schemata are termed *active knowledge* clusters in the CLCC model (see 4.3.10 below). The semantic context provided by active knowledge clusters propagates through the entire model: It is transferred directly to the stimulus attenuator, to act as the basis for the inclusion of stimuli by context relevance (a process termed semantic priming – Maxfield, 1997); and it is passed to working memory to allow for semantic based chunking, as well as provide meaning to the temporary constructions (partly leading to expectation based processing – Posner & Snyder, 1975). Because declarative memory contains almost exclusively learned information, it will be a source of a large degree of individual variation; however, this variation is not entirely idiosyncratic, as much semantic information is shared across populations such as cultural groups (Hirschfeld & Gelman, 1994) and professional groups (Goldman, 1986).

Inputs from: Media decoders and the folk-psychology and folk-physics modules, which provide perceptual inputs to be placed into wider semantic contexts.

Outputs to: The stimulus attenuator to implement the relevance bias; also to working memory, to ensure the formation of semantically coherent temporary structures. Finally, the system outputs to procedural memory, to give information about objects (hardness, weight, etc. as well as semantic information for speech and writing) which are required to plan and execute action in the context of the current semantic bias.

4.3.7 Procedural memory

Description: This node is responsible for storing sequences of behaviours (Squire *et al.*, 1993), including speech, walking, social responses such as greeting and maintaining eye contact, and simply conditioned reflex actions (Squire, 1999). It is included as a system separate to declarative memory in the CLCC model due to evidence that it is implemented by different brain circuits (Squire, 1999), and from clinical double-dissociation studies showing that lesions in particular brain regions negatively impacts declarative memory but not procedural memory, and vice-versa (Squire, 1999). Further evidence for the separation of these systems comes from Leeb *et al.* (2005) who showed that the *intention* to move a limb, even when actual movement is inhibited, generates a measurable EEG scalp potential. In the CLCC model, procedural memory has two roles: first, as suggested by Lee (2004), to initiate reflex reactions based on inputs from the folk psychology and folk physics modules, such as changing body posture (Freeman *et al.*, 2000), reaching for a virtual stimulus (Slater *et al.*, 1995c) or adjusting interpersonal distance in response to an agent (Bailenson *et al.*, 2001). The second role of declarative memory is the performing of goal-directed behaviours based on inputs from the folk physics and folk psychology modules. These behaviours include direct behaviours (such as reaching with an arm in response to an input to grab – see Waterworth & Waterworth, 2001) as well as symbolically mediated behaviours (such as clicking the mouse in response to an input to grab - Waterworth & Waterworth, 2001).

Inputs from: Temporary structures in working memory, which can act to inhibit as well as enhance the expression of behaviours based on the current semantic bias. A second input comes from the folk physics and folk psychology modules, to allow both goal-directed and automatic responses to objects. A final, indirect input is the set of active knowledge clusters in declarative memory, which allow for speech (by providing semantic information) as well as mediated behaviours (as when using an interface to mediate the VE).

Outputs to: The motor control system, for movement of the body or production of speech.

4.3.8 Motor control system

Description: This system contains the low-level neural and physiological controls for behaviour in the environment (including muscle movements and the motor aspects of speech). This node can be directly measured by means of physiological measures (as done by Meehan, 2001; see chapter 2 for a review of such measures). This system will also be of interest when considering behavioural presence measures (such as posture and sway – IJsselsteijn, 2004; see chapter 2).

Inputs from: Motor programs in procedural memory

Outputs to: Movement of the subject's body in the environment

4.3.9 Physical environment

Description: This is the physical environment, from which stimuli arise, and in which the subject interacts bodily. In many respects, this node represents embodiment, and underlines the importance of interaction in the world for presence (Lakoff & Johnson, 1999; Schubert *et al.*, 2002). With the exception of dreaming, where the reticular activating system circumvents normal perception (see 3.1.3), all stimuli arise from the physical environment, although, due to working memory limits, not all stimuli can be processed (Baddeley, 1986). Note that this is not the *mediated* or *virtual* environment (which has no explicit existence in this model – see 3.1.4 in chapter 3). This has two important implications: first, all stimuli (be they considered real or virtual) compete for processing on equal terms, without the benefit of semantic distinctions between them (these distinctions are only applied after a simulus has been selected for processing). Second, all motor movements made by the subject are expressed in the physical environment; movement in the virtual world occurs due to manipulation of the VE interface in the physical environment.

Inputs from: The motor control system, as well as from object to object interactions within the physical environment itself (which is a closed system in its own right).

Outputs to: Physical stimulation of the sensory organs which are converted into percepts by the sensory cortices, to enter at the stimulus attenuator.

4.3.10 Active Knowledge clusters

Description: These structures are highly transient, being created and discarded as the overall state of the model changes inside of a single experience. Although the CLCC model conceptualizes these clusters as objects, they consist of small, related clusters of active schemata in declarative memory (Rumelhart *et al.*, 1986; Rumelhart & Ortony, 1977). Due to the associative nature of declarative memory, active schemata automatically self-organize into meaningful clusters (as each concept is connected to semantically related concepts, allowing activation to spread between them - Rumelhart *et al.*, 1986). These clusters contribute significantly to the overall semantic relevance bias of the system, as they feed back to the stimulus attenuator (to filter out irrelevant stimuli), as well as to working memory, to provide semantic coherence to the current construction.

If active knowledge clusters stop receiving stimulation from working memory, their activation will gradually decay. One active cluster may also have its activation inhibited by another competing knowledge cluster, in which case it will decay more rapidly. (Rumelhart *et al.*, 1986; Rumelhart & Ortony, 1977). Regardless of how a knowledge cluster loses activation, it is not an instantaneous process. The more extensive the semantic network is which is activated, the longer it will take for the decay or inhibition to sweep across all connected schema (Rumelhart *et al.*, 1986). The amount of time taken for a cluster to lose its activation is referred to in the CLCC model as its *thematic inertia.* Clusters with higher thematic inertia are more likely to remain active and compete with other knowledge clusters, and are thus more capable of exerting a semantic bias over the system. Because thematic inertia is associated with the extensiveness of the particular schemata network, experts in particular content knowledge domains (who have more extensive knowledge networks – Ericsson & Lehmann, 1996) will have more thematic inertia for those content domains.

Inputs from: Bottom-up inputs from media decoders, the folk psychology module and the folk physics module. An active cluster can also pass activation (inhibitory or excitatory) to other, semantically similar clusters, via the networked structure of declarative memory.

Outputs to: Other knowledge clusters which are semantically related (both excitatory and inhibitory; see *inputs* above). The active knowledge clusters, as components of declarative memory, contribute to the overall semantic bias of the model (see 4.3.6 above).

4.3.11 Temporary working memory structures

Description: As with active knowledge clusters, these temporary structures are created and discarded in order to keep the overall system sensitive to changes in the environment. However, unlike the active knowledge clusters which represent activation patterns in fixed networks of knowledge, the temporary structures represent free-form clusters of data (Baddeley, 1986) which arise from the activation of declarative memory as well as from the use of working memory as temporary storage by the processing modules. They represent system state rather than structure; they represent all of the conscious information available to the subject (Rumelhart *et al.*, 1986). Due to the semantic bias transferred to working memory from declarative memory, these temporary structures are always constructed with a high degree of semantic coherence, and thus exert a great expectation bias on processing (Baddeley, 1986).

As these structures include both perceptual and semantic information, they are well suited to explaining content related effects in presence. In order for content to affect processing, the percept must first be decoded (requiring perceptual information) and then understood as meaningful objects (requiring semantic information). This is not the case with similar structures in other presence models, which incorporate very little or no semantic information (such as the SSMs of the MEC model – Wirth *et al.*, 2007). In the CLCC model, the set of temporary working memory structures is termed 'the construction of the environment' as these structures contain the conscious understanding which the subject has constructed by interacting in the environment (Bruner, 1990). It should be noted that although the temporary structures act to maintain semantic coherence, it is possible for an irrelevant stimulus to arrive from the stimulus attenuator demanding processing capacity (due to massive semantic dissimilarity). When this occurs, a sudden re-allocation of working memory occurs, and some temporary structures are discarded while others are changed drastically, leading to a new semantic interpretation of the scene – this is termed a *reconstruction*. Sudden reconstructions are associated with the break in presence experience, while gradual reconstructions (where working memory capacity is gradually allocated to tasks other than processing the environment) are associated with a *drift in presence* – see section 4.4.2 below.

Inputs from: Perceptual data arriving bottom-up from the stimulus attenuator, and conceptual data arriving top-down from the declarative memory system.

Outputs to: No direct output. These structures exist as data clusters in working memory, which are evaluated by the folk psychology module, the folk physics module or one or more media decoders.

4.4 Presence in the model

The CLCC model is strongly based on a general purpose cognitive model, as opposed to models which are specifically designed for explaining presence, such as those of Slater (2002), Biocca (2003), the FLS model of Waterworth and Waterworth (2001), and Wirth *et al.*(2007). Therefore, only the state of the model (and not the structure, which exists for processing in general) is pertinent when considering how presence occurs. In the CLCC model, presence is not a phenomenon of its own right, but rather side-effect of the model structure (as suggested by K. M. Lee, 2004). It is important to note that some of the structure of this model is inherent (such as the folk physics module, which is a species evolved trait), but other parts are acquired through learning (such as the media decoders). The state of the model is simply the distribution of activation among the structural components of the model. In this model, presence occurs mostly due to inference. At any moment, the system contains a set of active and semi-active knowledge clusters and behavioral scripts. Due to the associative nature of declarative memory, this pattern of activation allows contextualized inferences, and if any one inferential cluster receives enough activation to either lead to a temporary structure in working memory (a conscious thought) or to an overt behaviour, then one can say that that a particular cluster was 'selected' for expression (although the selection requires no more mechanism that the accumulation of activation and competition between active knowledge clusters).

The CLCC model implements presence as the perceptual illusion of non-mediation (Lombard & Ditton, 1997). This occurs when the subject is selecting a subset of stimuli (those encoding the virtual environment) from which to regulate their cognition and behaviour (this also owes something to the cognitive presence concept - Nunez & Blake, 2001). When a subject focuses on a VR display, processing resources will be allocated for those stimuli, and because the entire system is resource limited, resources will be taken from processing other stimuli and concepts (including the notion of 'this is a mediated display'). The more the subject interacts with the display, the more focused the semantic bias of the system, and the more likely it will be that resources will be allocated to processing it, simultaneously reducing the possibility of the subject processing the stimuli as being a display. Note however that although the illusion of non-mediation arises from focus on a set of mediated stimuli, the process is not only perceptual – it involves a complex interaction of top-down and bottom-up data.

4.4.1 Becoming present in an environment

Perhaps the best way to explain presence in this model is to illustrate how a subject becomes present in a VE, and how that presence experience might end. This example will begin with the case of a subject seated in from of a VR system, focusing some attention on the stimuli produced by the system (although this discussion will focus on becoming present in a VE, it can equally well apply to becoming present in any environment).

When the subject begins their experience, the sensory cortices will be processing stimuli reaching the sensory organs from various sources. However, the existing semantic bias of the system will tend towards excluding stimuli from the VE, or at the most, treating them as stimuli in the context of the larger situation (such as "noisy laboratory" or "office with computer"). However, if the stimuli are intense enough, some will be selected by the stimulus attenuator, and begin processing in working memory. Once there, some capacity will be allocated to creating a temporary structure for processing those stimuli. This will have require some of the capacity which was being used to process the larger scene to be withdrawn, effectively reducing the overall semantic bias of the system towards that scene. Due to the automatic chunking of data in working memory, if more stimuli are received from the VE which correlate with the first stimulus received, these will be chunked together into the same temporary structure.

Once the VE is represented in temporary structures in working memory, it will be evaluated by the folk physics and folk psychology modules. From this point on, the processing takes on a similar form to the MEC model's perceptual hypothesis testing phase (Wirth *et al.*, 2007), although the temporary structures are far more detailed than perceptual hypotheses, as they contain both spatial and semantic information. In the case of mediated environments, only some elements of the scene will be processed by one of the evolved modules, while all other elements will be handed off to one of the media decoders to process. If some immediate action is required by either the folk physics module (such as ducking away from a fast moving object) or the folk psychology module (such as smiling in response to a smile), then these modules can initiate this behaviour by stimulating procedural memory directly. Otherwise, they will pass the decoded stimuli to declarative memory, to attach semantic meaning in the context of the current semantic bias. Stimuli which are passed through to the media decoders are processed to extract semantically meaningful objects, but media decoders, (as learned structures) cannot directly lead to automatic or reflex actions.

Temporary structures in working memory then begin to activate knowledge clusters in declarative memory. The activation is automatically passed to semantically related concepts, and simultaneously, the inhibition of unrelated clusters. If some clusters are already partly activate (due to the previous system state or thematic inertia), they will tend to dominate even with slight activation. Also, if the content of the VE is one for which the subject is an expert, knowledge cluster activation will be more extensive, leading to larger semantic effects. The most active knowledge clusters in memory will impose a semantic bias which is passed on to the rest of the system – in particular, it is passed down to working memory, where it leads to a bias towards forming temporary structures which are semantically related to the current bias, and to the stimulus attenuator, which tends to exclude stimuli not relevant to the current bias.

The final path of activation in the model is from the active knowledge clusters to procedural memory. Generally speaking, no single path leads to the activation of a particular action script; rather, each cluster partly activates a number of bias-appropriate scripts, which exist as possible responses to the current situation. A number of active knowledge clusters need to converge on a single script before it is executed. When the subject does engage in some action, the changes in the environment they produce are encoded by the sensory cortices, and that becomes an input to be evaluated by the stimulus attenuator for inclusion in further processing.

Note that a powerful feedback loop exists between declarative memory and the stimulus attenuator and working memory, which acts to set and hold a semantic bias. This integration between different levels of abstraction (reminiscent of the levels of processing model of Craik & Lockhart, 1972) is similar to the integration which the LOP model requires to achieve optimal focused presence (Riva *et al.*, 2004). Like the LOP model, the CLCC model achieves the greatest sense of presence when all levels are processing a coherent stimulus set. In effect, the semantic bias in the model represents the interpretation or construction of the situation; and each construction will attempt to stay active by keeping out irrelevant stimuli (partly due to the cost involved in reconstruction – see below). This process is similar to the hypothesis testing process in the MEC model (Wirth *et al.*, 2007) – although instead of collecting positive evidence, the CLCC model ensures coherence by a simple excitation/inhibition competition mechanism among the active knowledge clusters in declarative memory. This means that relevant perceptual stimuli will further activate the current active knowledge clusters, leading to a stronger, more effective bias. Therefore, if a subject begins to process the VR system, and there is no competition from semantically unrelated stimuli, the VE will eventually become the dominant bias in their processing, which will lead to VE relevant inferences and actions favoring the semantic meaning provided by the VE, which can be considered a state of high presence (Nunez & Blake, 2001).

4.4.2 Failures of presence: Breaks and drifts

A unique feature of the CLCC model is its inclusion of both a stimulus attenuator and working memory allows the CLCC model to explain presence as a binary and a graded experience. The binary character of presence (as occurs during a break in presence; Slater & Steed, 2000) can be explained as occurring at the stimulus attenuator level – recall that stimuli which widely violate the semantic bias of the system are allowed to enter processing (to keep the model adaptable to changing environments). The large degree of mismatch between the new stimulus and the current construction will triggers a rapid reconstruction of the environment (a break in presence), which will effectively end the presence experience.

However, presence can also fail due to a number of other reasons – preoccupations (Slater & Steed, 2000), a competing second task (Freeman *et al.*, 2000), or a simple case of the subject becoming bored with the content (Wirth *et al.*, 2007). Such situations are associated with a gentle decrease of the presence experience, in what might be termed a ‘drift out of presence’. Under these conditions, the subject has begun to allocate working memory to processes other than cognition in the VE (working out the monthly grocery budget, the mental gymnastics required by the second task, etc.). As less working memory is allocated to processing the VE, a less detailed construction of the VE will result. This in turn gives a reduced semantic bias, fewer inferences about the VE, and a lower likelihood of VE-relevant action. If this competition continues, it may lead the subject to stop being present altogether (this could occur when no working memory is assigned to processing the VE, such as when sensory stimuli are missing, or when working memory is overloaded with other tasks). Evidence for such a drift phenomenon could be taken from an interference or simultaneous loading task study (as used by Baddeley, 1986). Unfortunately, most of the studies which look at the effects of task performance on presence (for instance, Bystrom & Barfield, 1999) do not manipulate task difficulty, so there is no direct evidence for this claim (see Welch, 1999 for an elaboration of this argument).

4.5 The immersion-presence relationship

An important general finding in the literature is that higher degrees of immersion lead to more intense presence experiences (see section 3.3.1.3 in chapter 3 for a review of some of the evidence). Predicting this relationship is arguably the most basic validity requirement for any presence model, given the amount of evidence supporting the relationship. The CLCC model is able to explain the immersion-presence relationship in a way similar to other models – by considering how attention and resources are allocated (Biocca, 2003; Slater, 2002; Waterworth & Waterworth, 2001; Wirth *et al.*, 2007). More immersive systems effectively decrease the number of stimuli which encode non-VE attributes (Slater, 1999). This affects multiple levels of the CLCC model. First, because there are fewer stimuli competing with the existing semantic bias, more working memory will be available for processing the VE. Second, the effective amount of capacity available for processing the VE is increased by virtue of the fact that semantically related material can be chunked together (Baddeley, 1986); in an immersive system, most of the information arises from the same virtual objects and will therefore be highly semantically consonant. Third, the existence of fewer non-VE related temporary structures implies a decreased probability that the model will require a reconstruction to deal with an unexpected stimulus. The dominance of VE related temporary structures in working memory will lead to an increase in the number of VE related active knowledge clusters active in declarative memory, and thus an increase in the overall VE-related semantic bias. This translates to increased expectations related to the VE, more inferences about the environment which are semantically consistent with the content of the VE, and therefore a more intense presence experience.

4.6 Forms of presence beyond spatial

A number of models restrict themselves to explaining only spatial forms of presence (such as the environment selection and MEC models of presence – Slater, 2002; Wirth *et al.*, 2007). However, as argued in section 2.2 of chapter 2, many researchers who define or measure presence consider it to be a complex, multi-factor construct (see for instance Lessiter *et al.*, 2001; Lombard & Ditton, 2004; Witmer *et al.*, 2005). The CLCC model allows the explanation of both spatial presence, and other forms of presence such as engagement and naturalness (Lessiter *et al.*, 2001; Lombard & Ditton, 2004). It should be noted, as outlined in section 2.6, that the current form of the CLCC model is, for reasons of expediency, a model of individual forms of presence, and although it may hold true for social forms of presence, has not been validated for those phenomena.

Engagement, the sense of psychological involvement with the VE content (Lessiter *et al.*, 2001) can be explained by considering semantic networks in declarative memory. When processing a scene, knowledge clusters which are semantically related to the scene will become active and spread their activation to similar clusters. The more extensive these networks, the more thoughts and experiences will be associated by the subject with the scene, which will give a greater sense of personal connection to the scene. The model reflects the suggestion by Wirth *et al.* (2007), that engagement with an environment is a function of higher-order cognitive involvement. It should be noted however, that enjoyment is not a necessary consequence of this process (as suggested by Lessiter *et al.*, 2001) – if the subject's experiences with the content are negative, then the experience can be a negative one – evidence of this comes from

Rothbaum *et al*. (1999; , 2001), who showed that subjects suffering from post-traumatic stress disorder experienced anxiety and negative affect when experiencing a VE of the situation which brought about their trauma; and by Robillard (2003), who showed that subjects suffering from a phobia experienced both increased PQ scores and anxiety when entering a VE containing the object of their phobia.

Naturalness, the sense that the content of the VE is lifelike or realistic (Lessiter *et al.*, 2001), operates similarly to engagement. Unlike engagement, which is an automatic reaction to content, naturalness requires an implicit comparison of the content of the scene with previous knowledge of the content to determine its realness. This comparison is done automatically by the CLCC model. Recall that when activation spreads in declarative memory, knowledge clusters are activated to varying degrees. Sometimes a knowledge cluster is activated, but not enough to be expressed as conscious in a temporary structure in working memory. This condition represents an implicit expectation, which can be matched (i.e. activated to consciousness) or not matched by incoming sensory data from the stimulus attenuator (similar to extended presence - Riva & Waterworth, 2003). A large number of these unmatched expectations in the model will give the subject a sense that the scene is unfamiliar or unusual. However, if the scene matches these expectations, they will have a sense that the scene is complete, familiar and realistic.

4.7 How the model deals with the five problems

To demonstrate the theoretical validity of the CLCC model relative to the models reviewed in chapter 3, the following section will examine how the CLCC model explains each of the five presence problems used in section 3.1 in chapter 3.

4.7.1 The book problem

Presence in the CLCC model is characterized by temporary structures in working memory devoted to processing the VE, and a model-wide semantic bias exerted from declarative memory. This state can be equally achieved either from information from the folk physics (or folk psychology) modules, or the media decoders. As media decoders are learned structures, this implies that presence can arise practically from any information encoding source, including books. This explains the first part of the book problem – how books (or any low immersion medium) can lead to presence. The second part of the book problem is why books lead to consistently less intense presence experiences. This can be understood as part of the presence-immersion relationship described in 4.5 above. Unlike the folk physics and folk psychology modules (which have dedicated neural circuits for processing) media decoders need a particular amount of working memory for decoding. This means that less working memory capacity will be available for temporary structures, and therefore a reduced semantic bias in favour of the VE, resulting in reduced presence.

4.7.2 The physical reality problem

For this problem, one must examining how working memory is allocated, and how the semantic bias affects the stimulus attenuator. When someone is lost in their thoughts, they are involved mostly in top-down processing (Biocca, 2003). As declarative memory feeds directly down into working memory, it is possible to have temporary structures formed only from semantic information. Thus, two possible scenarios exist for someone not processing the outside environment: The first is if the subject is engaged in semantic processing that requires so much capacity that none is left to allocate to perceptual stimuli arriving from the stimulus attenuator (as would occur with a particularly difficult task). In the second scenario, the subject is not engaged in a difficult task, but the semantic content which is being processed is associated with a vast network of knowledge structures in declarative memory (as would occur when someone is lost in recalling personal experiences, or lost in a paracosm of their own making – Cohen & MacKeith, 1991). Under these conditions, the high degree of knowledge structure activation leads to an unusually specific semantic bias fed down to the stimulus attenuator, such that almost all incoming stimuli are recognized as irrelevant, and thus filtered out.

4.7.3 The dream state problem

In the CLCC model, the presence experiences during dreaming can be explained in a similar way to the physical reality problem – temporary structures are formed in working memory using only top-down activation from declarative memory. The difference is that unlike during waking presence, activation of the knowledge clusters does not come from external stimulation or by willful act (both of which do not occur while dreaming); rather the activation of knowledge clusters during dreaming comes from random activation of declarative memory which occurs naturally during REM sleep (Hobson *et al.*, 2000). Once this random activation has led to the creation of active knowledge clusters, this can be fed down towards working memory to form temporary structures and a semantic bias in the model.

4.7.4 The virtual stimuli problem

This problem poses that virtual stimuli (stimuli which arise from a virtual environment) and 'real' stimuli are in fact the same, as they all arrive as physical energy at the subject's sensory organs without information of their origin. The problem is how a subject can select a subset of these stimuli and process them as if they compose a coherent environment. The CLCC model is capable of explaining the virtual stimuli problem completely. Stimuli are selected for further processing by the stimulus attenuator, and they will be added to one or more temporary structures in working memory. It is only once the stimuli are adopted into a temporary structure that they acquire semantic meaning (such as "real" or "virtual"). Note that as a particular semantic bias takes over the model, expectation-based processing increases, and all stimuli are more likely to be interpreted as belonging to the semantic construction, and incongruent stimuli are less likely to be allocated working memory capacity. Therefore, as long as the VE display contains semantic coherence, it will be incorporated into a small set of temporary structures, and will be experienced as a coherent environment, discreet from other environments which may have their own constructions encoded by other temporary structures.

4.7.5 The inverse-presence problem

The essence of this problem is that the subject experiences a real experience as mediated. As discussed in chapter 3, none of the extant presence models are able to explain this problem adequately, making it a significant theoretical hurdle. An important clue to solving this problem is noting when it occurs – usually under conditions of extreme violence, beauty or drama (Timmins & Lombard, 2003). These are situations which most subjects have probably only experienced through some mediation technology (such as film or television). If it is indeed the case that subjects experience inverse presence when encountering situations with which they have more mediated than unmediated experience, then the CLCC model is well able to explain the phenomenon. Consider an example of a subject experiencing a car accident, something which they have experienced only in films and television programs. Recall that semantic meaning is derived from previous experience encoded in declarative memory. As the car accident is more closely associated with mediation, the experience itself will be closely semantically associated in declarative memory with features of mediation (such as the sense of being an observer rather than an actor) and therefore create expectations associated with the mediated experience (such as waiting for dramatic music to play). A subject who is in a VE experiencing something ordinary can feel present because their expectations and semantic associations are linked to other non-mediated experiences. The unfortunate subject in the car crash experiences the opposite - a real event triggers expectations and semantic associations which are linked to mediated events. It is important to note that in the CLCC model, no experience or event is tagged as ‘real’ or ‘virtual’ – this occurs implicitly by the associative nature of declarative memory. If an event that is normally associated with non-mediation occurs due to stimuli from a mediated source, then this is a potential source of presence; whereas if an event that is normally associated with mediation occurs, then this is a potential source of inverse presence.

4.8 Conclusion

This chapter has detailed the capacity limited, cognitive constructionist model of presence, in which a subject constructs their own experience of an environment based on perceptual and semantic data, while being constrained by the limits imposed by working memory. This chapter has argued that the CLCC model can explain Lombard & Ditton’s (1998) concept of presence and that it can predict forms of presence beyond spatial (such as engagement and naturalness). Furthermore, the model is unique in that it can, by combining working memory with an attention filter, explain presence as a continuous and binary phenomenon simultaneously.

The theoretical validity of the model was evaluated in two forms: First, by considering the model’s ability to explain the well-established presence-immersion relationship; and second, from the model’s ability to explain the five theoretical problems in presence (which none of the models surveyed in chapter 3 was able to do satisfactorily). Although the CLCC model shows a high degree of theoretical coherence, some important aspects require empirical validation. In particular, the role of limited capacity in the model, and the importance of semantic coherence for the presence experience have not been thoroughly investigated in the literature. The following part of the dissertation will present a set of empirical studies which aim to provide evidence for the validity of the CLCC model as a whole.

Part II:

Empirical evaluation of the CLCC model

This part of the dissertation presents six empirical studies (using data from 1080 subjects) which were designed and conducted to evaluate specific aspects of the CLCC model. Studies 1, 2 and 3 (in Chapters 5, 6, and 7 respectively) examined the role of working memory and attention (capacity limits) in the presence experience in laboratory studies. Studies 4 and 5 (in Chapters 8 and 9) are two on-line surveys conducted on computer gamers (as expert user populations) to determine the role of cognitive styles and content expectation on presence. Finally, Study 6 (in Chapter 10) presents a laboratory experiment which manipulated content and non-diegetic music to examine the role of semantic coherence in presence.

Chapter 5: Study 1 - Effects of working memory load on presence

A major component of the CLCC model of presence is the working memory module, and therefore working memory effects (and the associated capacity limits effects) are of interest. As detailed in 4.3.3 in chapter 4, the inclusion of working memory in a model of presence is a novel theoretical contribution, and there is thus no empirical data in the literature to support that aspect of the model. This chapter describes an experiment to evaluate the role of working memory on presence, by adapting the working memory dual task paradigm developed by Baddeley (1986) to run in a virtual environment.

5.1 The working memory dual task paradigm

Baddeley (1986) developed a simple experimental paradigm for investigating if working memory plays a role in some particular task, and for uncovering which of the two working memory systems is involved in the task. The logic behind the paradigm is simple: if a task uses working memory, performance on that task should degrade if not enough working memory is available to complete it; and because the amount of working memory available is finite, one can create a condition of reduced working memory by giving a subject a concurrent interfering task. In a classic example, Baddeley and Lieberman (1980) had subjects remembering a list of words using either a verbal or a spatial mnemonic, while simultaneously performing a spatial task. The results showed that those subjects using the spatial mnemonic had poorer performance than those using the verbal mnemonic, indicating that spatial and verbal working memory were separate systems. By manipulating the degree of working memory load provided by the loading task (such as by altering the length of the number to be recalled), it is possible to examine how much working memory a particular task requires (Baddeley, 1998).

The present study used this paradigm: subjects are placed into one of two conditions (spatial loading task or verbal loading task), and then given one of four different levels of working memory load. This design allows the exploration of whether presence makes use of verbal or spatial working memory (or both or none), and also the degree of working memory required for presence in the given environment.

5.2 Predictions about working memory made by the CLCC model

As this study is specifically designed as a test of a component of the CLCC model, it is necessary to provide a set of predictions made by the CLCC model for a dual task situation where one of the tasks is experiencing presence. The following predictions can be made:

1. *Performance on the main task will degrade as a function of loading task difficulty* – The traditional loading task used in this paradigm is a list recall task. Because a longer list requires more working memory to store, longer lists are more likely to have recall errors than shorter ones (this is referred to as the word length effect – Baddeley *et al.*, 2002). This is predicted by the working memory nodes of the CLCC model, but this effect is expected whether the subject is experiencing presence or not; it should be considered as a secondary prediction, not central to providing evidence for the CLCC model.

2. *Working memory load will negatively affect presence* – The CLCC model proposes that presence occurs when the construction of the virtual environment becomes so extensive (in terms of temporary structures and active knowledge clusters), that it dominates as the basis for inference and behaviour. In order for this to occur, temporary structures must be created in working memory, to provide a coherent bridge between top-down and bottom-up data. Furthermore, the CLCC model predicts that the more temporary structures dedicated to processing the VE are present in working memory, the more intense the presence experience will be, as they will allow for the activation of more knowledge clusters, and therefore a wider range of VE-consistent inferences. Therefore, the CLCC model predicts that loading a subject with a working memory loading task while experiencing an environment will lead to a presence experience which is impoverished as a function of the load imposed by the second task.

3. *Working memory load will affect cognitively higher forms of presence (such as engagement) more than spatial presence* – The CLCC model proposes that presence occurs partly as a consequence of activation spreading between temporary structures in working memory and active knowledge clusters in declarative memory, in order to provide semantic meaning. According to the model, spatial presence occurs due to low-level cognitive processes which decode the sensory stimuli and create a mental model of the space in temporary structures in working memory. However, the existence of these temporary structures leads to activation of semantic information in declarative memory which feeds back down to working memory to create further temporary structures, which encode progressively more semantic information about the environment. Therefore, the CLCC model predicts that spatial presence requires less working memory than the higher cognitive forms of presence such as engagement and naturalness. This means that when the subject is faced with a working memory loading task, they will experience a greater reduction in their engagement and the naturalness of the experience than in their spatial presence.

4. *Loading the visual WM system should impact spatial presence more than cognitively higher forms of presence (such as engagement), but loading the verbal WM system should affect cognitively higher forms of presence more than spatial presence* – This prediction follows prediction 3 above. Because semantic information is modeled as playing a larger role in engagement and naturalness than in spatial presence, it is predicted that these forms of presence will make more use of verbal working memory (which is associated with semantic processing – Carpenter & Just, 1989); conversely, because spatial presence deals more with creating a spatial model primarily from visual information, it is predicted that spatial presence will make more use of spatial working memory. The CLCC model therefore predicts that a spatial loading task will negatively affect spatial presence more that than engagement or naturalness; while a verbal loading task will more negatively affect engagement and naturalness than spatial presence.

5.3 Sample

Undergraduate psychology students were offered course credit for participating. A total of 177 students participated - 141 women and 36 men. The mean age of the subjects was 21.3 years (S = 3.23). The sample was asked to self-rate their computer experience, game playing experience, and knowledge of virtual reality, each on a 6 point scale (0 = no knowledge/experience, 5 = expert) as well as video game playing frequency (0 = never play, 5 = play once a day). The sample can be considered as novices in all these variables (see Table 5.1 below).

	M	*SD*
Computer experience	1.615	0.592
Game playing experience	0.203	0.807
Game playing frequency	0.740	0.746
Knowledge of VR	0.316	0.585

Table 5.1: Sample expertise and experience of VR technology

5.4 Apparatus

The study ran in a dedicated room with five desktop computers, each with the same hardware configuration (see Table 5.2 below). This produced a measured update rate on the environments which ranged between 17Hz and 28Hz. The room was made dark for the duration of the study. The computers were separated by partitions such that each subject could only see their own display.

Hardware Used	
Display:	17” Samsung Syncmaster 750 CRT
Graphics card:	GeForce 6200, 128MB RAM
Processor:	Intel Pentium 4, 2.8GHz
RAM:	512MB, DDR333
Input devices:	Keyboard and optical mouse
Sound:	Stereo, by headphone

Table 5.2: Hardware specification of the desktop machines used

5.4.1 Virtual environments

Two VEs were used in the study – a training VE used to familiarize subjects with the system and mode of interaction, and a main VE in which the study was conducted. The training VE (see Figure 5.1 below) consisted of eight rooms spread over two levels, and contained two panels and locked doors (see 5.6 below) for the subjects to practice with. The main VE represents a modern hospital spread over 4 floors. Apart from the subject, the VE is deserted (see figures 5.2 and 5.3 below).

The VEs were rendered using the Genesis3D engine (http://www.genesis3d.com), at a resolution of 1024x768x32. The system also presented spatialized environmental audio (through stereo headphones), and the subject could hear their footsteps in the VE as they walked. Control in the VE was by means of the keyboard and mouse, using the Quake Keys control method (Dalgarno & Scott, 2000).

5.5 Procedure

The experiment ran over a period of four weeks. Subjects arrived at the study venue, and the researcher explained that the study was looking at psychological aspects of virtual reality. Subjects were then randomly assigned to one of the eight experimental conditions (see 5.6 below for the design). Subjects were given an introductory questionnaire measuring biographical information, and their level of experience with VR related technology.

They were then introduced to the training VE and shown how to use the controls and given practice on the experimental task (see 5.6 below).Once the researcher had determined that the subjects were proficient in navigating around the VE and in performing the experimental task, the training session was ended – this took no more than five minutes.

Figure 5.1: Image from the training VE – the grey door on the left is to practice the door opening task.

Figure 5.2: Image from the hospital environment

Figure 5.3: Image from the hospital environment

The subjects were then given a brief scenario on paper for their experience (see appendix E for the scenario text) – they were told that they were construction workers accidentally locked on the top floor of a hospital they were working at, and their task was to proceed to the bottom floor of the hospital building, where they could leave the building. The subjects then donned headphones, and the main VE was started. The subjects performed the experimental task in the VE for 15 minutes. After this, the subjects were given the ITC-SOPI, followed by a measure of their working memory capacity (see measures below for a detailed description of measures). As the measures of working memory used are stable and immune to manipulation effects (Baddeley, 2000), it was not necessary to counterbalance the order of the measures.

5.6 Experimental task

The subjects were given a working memory loading task (rehearsal and recall of lists of random strings of information) to perform during the VE experience, based on the tasks used by Baddeley *et al.* (Baddeley, 1986; Baddeley *et al.*, 2002). Two versions of the task were created, one for each of the two working memory systems – one loaded the spatial system (the visuospatial sketchpad), and the other loaded the verbal system (the phonological loop). Each of these tasks was crossed with four different levels of load (3, 4, 5 or 6 chunks), to create a 2x4 factorial design. Each subject was randomly placed into one of the eight resulting conditions (see Table 5.3 below for the number of subjects used in each condition). Both factors (WM load and WM system loaded) are between-subject factors, with each subject contributing data to only one of the eight cells.

WM system loaded	***WM load (chunks required to complete task)***			
	3	***4***	***5***	***6***
Verbal (phonological loop)	23	24	26	20
Spatial (visuospatial sketchpad)	20	20	24	20

Table 5.3: Study design and number of subjects randomly assigned into each condition

The task was to navigate through the environment by opening locked doors. The VE was created so that only one particular path led from the starting point on the top floor to the exit in the basement level. Along this path, 15 locked doors were placed. The doors could be opened by typing in a code sequence (randomly determined, and its length corresponded to the load level of the task). The codes could be found by accessing panels placed in the VE at some distance from the doors. The subjects thus had to search for the panels; obtain the code, and rehearse it in working memory while they navigated to the door, and then use the code to open the door. If the subject had forgotten the code when they reached the door, they had to return to the panel to get the code again. The VE was structured so that there was only ever one available panel for the next locked door – this way, subjects only had to rehearse a single code at a time. Panel access and code input was done by clicking on on-screen buttons with the mouse; subjects stopped next to either a panel or door in the VE, and pressed the left mouse button to being up the panel or door interface.

The lock and panels used either number strings (verbal load condition) or a grid of unmarked buttons (spatial load condition). Each lock type is analogous to standard working memory procedures used by Baddeley (Baddeley, 1986; Baddeley *et al.*, 2002) as working memory loading tasks. In the *digit span* task, the subject is given random digit strings which they must repeat after a delay; the load of the task is the length of the string. In the *block tapping task*, the subject is shown an array of similar cubes, and a sequence is tapped out on them. The subject must then repeat sequence by tapping the blocks after a delay; the load of the task is the length of the tapping sequence (Baddeley, 1986; Nelson *et al.*, 2000; Vandierendonck *et al.*, 2004).

Figure 5.4: Panel providing the verbal code (length of 3)

The VE system recorded the number of successful and failed attempts at opening each door, as a measure of task performance. In the case of the verbal condition, a different font was used in the panel and the lock; this ensured that subjects could not use the shape of the numbers (and thus make use of spatial working memory) to recall the code. No digit was used more than once in a code. Figure 5.4 below shows the verbal condition panel interface, and Figure 5.5 shows the verbal condition door interface.

In the spatial load condition, a square grid of 3x3 buttons was used. In the panel, the code sequence was played back (each button was illuminated for one second, with a 250 millisecond delay between buttons), and the subject could watch the sequence as many times as they wished by clicking the 'show' button on the panel. No block was used more than once in the code sequence. In the door interface, the subject had to click the sequence in order to open the door – buttons remained lit as the subject clicked the code. Figure 5.6 shows the spatial load condition panel interface, while Figure 5.7 shows the spatial condition door interface.

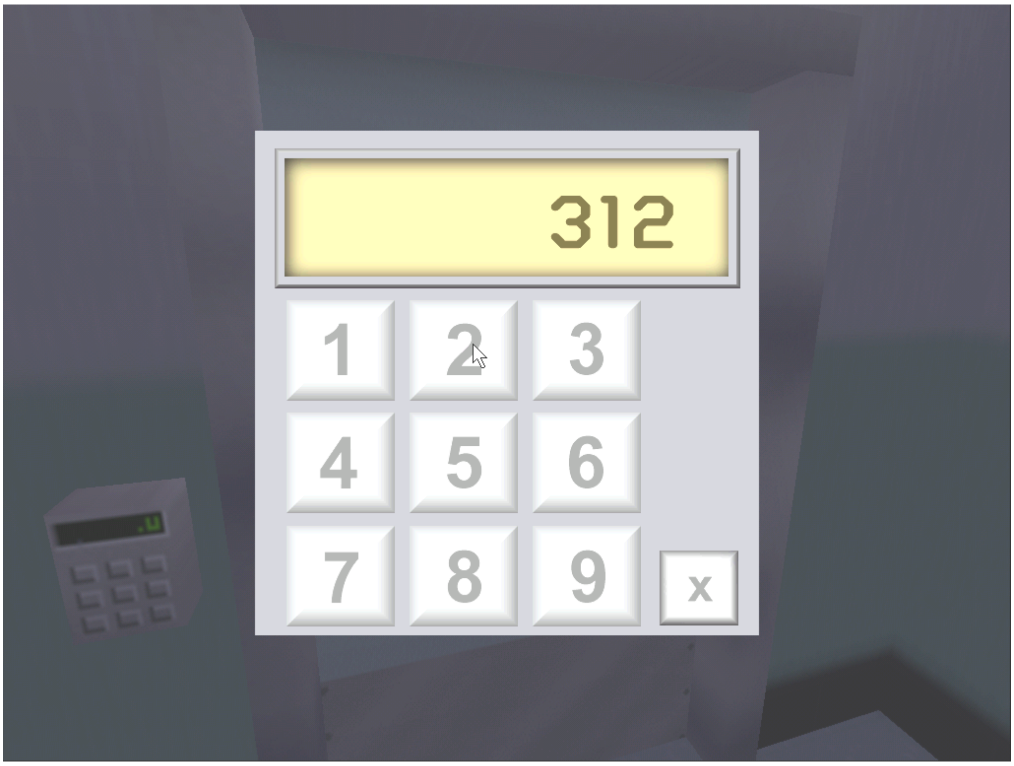


Figure 5.5: Lock for the verbal code (note a different font is used from the panel)

5.7 Measures

The main dependant variable (presence) was measured using the ITC Sense of Presence Inventory (ITC-SOPI - Lessiter *et al.*, 2001). The choice of this measure is discussed in 2.6 in chapter 2. This questionnaire measures four factors of the presence experience: Spatial presence, engagement, naturalness and negative effects. These are defined as follows (from highest to lowest degree of variance explained):

> *Spatial presence:* A sense of physically being in the VE, and of interaction with the objects in the VE.
>
> *Engagement:* A sense of psychological involvement and a tendency to enjoy the VE experience.
>
> *Naturalness:* A sense that the VE is believable and lifelike, or realistic.

Negative effects: Negative physiological reactions to the VE experience such as dizziness, eyestrain and headaches.

The ITC-SOPI is a particularly useful measure in that each of the factors provides a separate score for the experience. It is thus possible to independently examine effects on particular aspects of the presence experience.

As a control for individual differences in working memory capacity (Cowan, 2001), subjects were measured with two measures of working memory capacity: The digit span task with immediate recall (Waters & Caplan, 2003) which is a measure of verbal working memory capacity, and the Corsi block-tapping test with immediate recall (Vandierendonck *et al.*, 2004) which is a measure of spatial working memory capacity. Both of these measures are considered to be valid and reliable for experimental use (Moye, 1997; Waters & Caplan, 2003). The digit span task presents subjects with a string of digits, which they must rehearse over a short delay and then repeat. The length of the string begins at three, and is increased by one digit if the string is correctly repeated. If the subject makes an error, a new string of the same length is given. When the subject makes two errors in a row, the test is complete – the length of that string is the subject's verbal working memory capacity.

The Corsi block tapping test is similar: A number of blocks is placed before the subject, and a sequence is tapped out on the blocks. The subject must then repeat the sequence. The length of the sequence increases on successful repetition, and the test completes when two errors in a row are made. All measures were implemented as computer-based tests.

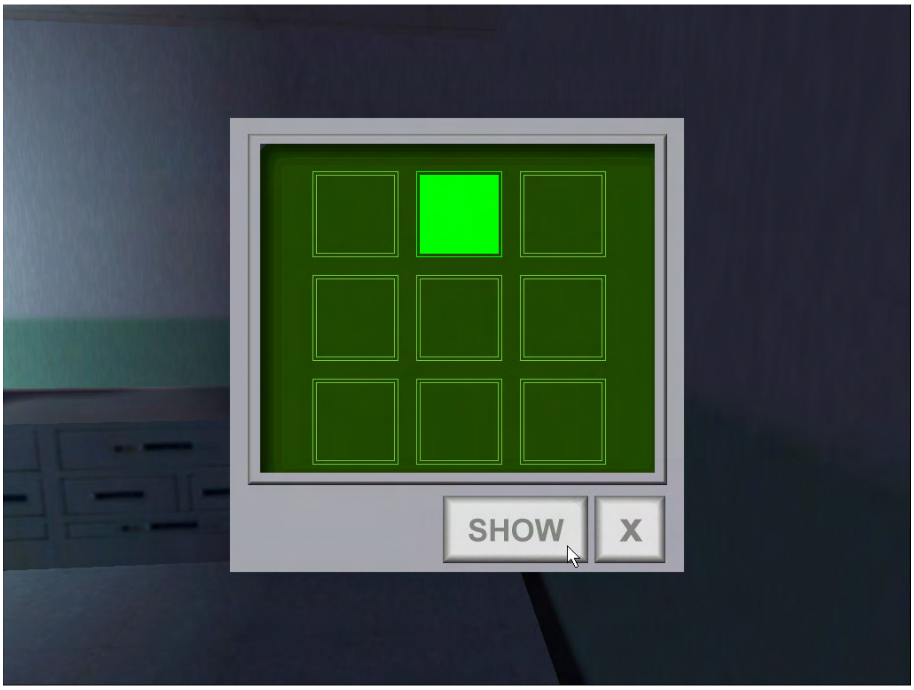


Figure 5.6: The panel showing the spatial code. The sequence of blocks is shown illuminated, and the sequence can be repeated using the 'show' button.

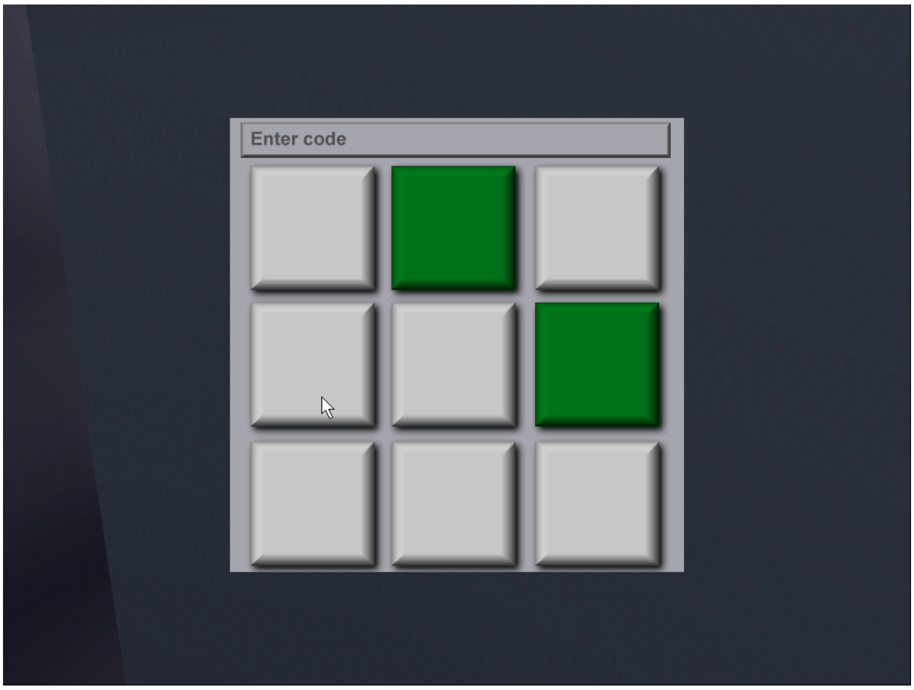


Figure 5.7: The lock for the spatial load condition. The green blocks have already been clicked by the subject.

5.8 Analysis & results

5.8.1 Effectiveness of randomization

There was no difference among the eight conditions on spatial working memory capacity ($F(7, 169) = 0.972$, $p < 0.453$) or verbal working memory capacity ($F(7, 169) = 0.819$, $p < 0.572$). Similarly, there were no differences between the conditions in terms of computer experience ($F(7, 169) = 1.768$, $p < 0.093$), game playing experience ($F(7, 169) = 1.049$, $p < 0.398$), game playing frequency ($F(7, 169) = 1.154$, $p < 0.332$) and knowledge of VR ($F(7, 169) = 1.307$, $p < 0.250$). The number of subjects in each condition (shown in Figure 5.3 above) indicates an even distribution of subjects ($\chi^2 = 0.195$, $p < 0.978$). Although there are a disproportionately high number of women subjects (79%), a contingency table analysis shows that the male subjects were evenly distributed among conditions ($\chi^2 = 0.198$, $df = 3$, $p < 0.978$).

5.8.2 Evaluation of experimental manipulation

To establish that the working memory manipulation was successful, an examination of task performance (number of doors opened during the 15 minute exposure) was conducted. A one-way analysis of variance using the WM load condition as independent variable and task performance (number of doors opened) as dependent variable shows no significant difference between load conditions ($F(3, 173) = 2.208$, $p < 0.08$), although the means profile (Figure 5.8 below) shows a downward linear trend which may indicate that the task difficulty manipulation was partly successful.

The lack of difference among conditions may have been attributable to individual differences in working memory capacity. To control for this, *effective working memory load* was calculated for each subject. This is defined as the experimentally imposed load divided by the subject's measured working memory capacity. Note that effective memory load is dependent only on the load condition the subject was placed in, and the non-loaded portion of working memory would have been available to use in presence processing. The average effective working memory load for the spatial load condition was 0.850 (85% of WM loaded by the task), while for the verbal load condition it was 0.517 (51.7% of WM loaded by the task). The difference between these is significant ($F(1, 174) = 94$, $p < 0.0001$), indicating that the verbal condition task was easier than the spatial condition task. Using this effective load as a predictor of task performance in a linear regression gives a significant model ($F(1, 174) = 39.706$, $p < 0.0001$, $R^2 = 0.185$). This indicates that controlling for individual differences in working memory capacity, the manipulation of task difficulty was successful, although the manipulation was weaker than one would have liked – ideally, the manipulation would have been strong enough to interfere with task performance irrespective of individual differences.

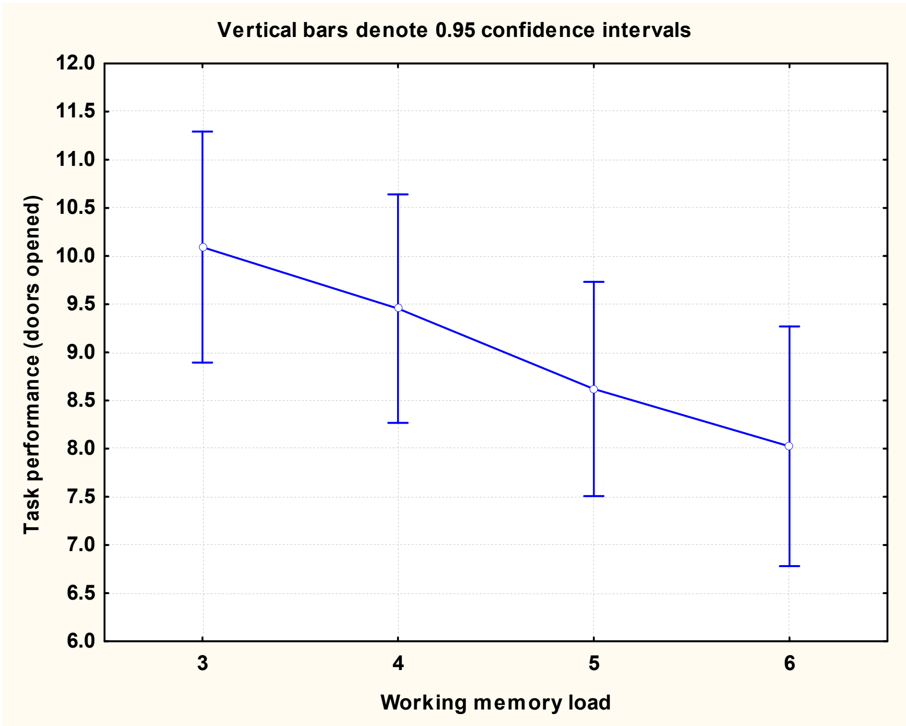


Figure 5.8: Task performance as a function of working memory load, showing the effect of the experimental manipulation (note that the ANOVA result is not significant: p < 0.08)

5.8.3 Effect of working memory load on presence

The data were analyzed with a set of general linear models (GLM), which is similar to a multiple regression analysis, but allows the modeling of a single continuous dependent variable using a mix of continuous and categorical predictors (Hastie & Tibshirani, 1997; Neter *et al.*, 1988). Each of the ITC-SOPI factors was modeled using the working memory system being loaded (spatial and verbal), and the effective working memory load as predictors. This analysis was preferred over using a 2-way analysis of variance using working memory system loaded and manipulated working memory load as factors, because as argued in 5.8.2 above, effective working memory load is a more sensitive measure of the working memory manipulation than the load condition.

5.8.3.1 ITC-SOPI spatial factor

The Cronbach's alpha for this factor was 0.93. For this factor, the main model is not significant: $F(2, 173) = 0.439$, $p < 0.645$. See Table 5.4 below for partial effects (the intercept results have been excluded from the table for the sake of clarity).

Effect	df	F	*p*
Effective WM load	1	0.639	0.425
WM system	1	0.006	0.937

Table 5.4: GLM results for the ITC-SOPI spatial factor. Intercept has been excluded for clarity.

5.8.3.2 ITC-SOPI engagement factor

The Cronbach's alpha for this factor was 0.89. For this factor, the overall model is significant: $F(2, 173) = 3.876$, $p < 0.022$, $R^2 = 0.041$. See Table 5.5 below for partial effects (the intercept has been excluded from the table for the sake of clarity). The analysis of predictors shows that WM system is the only significant predictor, showing a higher average score for the spatial load condition (M=3.438) than for the verbal condition (M=3.102). This difference of 0.336 points is modest if taken in the context of scores obtained for the engagement factor as a whole: they ranged between 1.461 and 5.000, with a mean score (across all conditions) of 3.270.

Effect	df	F	*p*
Effective WM load	1	0.812	0.369
WM system	***1***	***7.054***	***0.008***

Table 5.5: GLM results for the ITC-SOPI engagement factor (significant predictors highlighted). Intercept has been excluded for clarity.

5.8.3.3 ITC-SOPI naturalness factor

The Cronbach's alpha for this factor was 0.82. For this factor the overall model is again significant: $F(2, 173) = 3.208$, $p < 0.043$, $R^2 = 0.035$. See Table 5.6 below for partial effects (the intercept results have been excluded from the table for the sake of clarity). The analysis of predictors shows that WM system loaded is the only significant predictor, showing a higher average score for the verbal condition (M=3.267) than for the spatial load condition (M=2.938). As with the engagement difference shown above, this difference of 0.329 points is modest if taken in the context of scores obtained for the naturalness factor: they ranged between 1.000 and 5.000, with a mean score (across all conditions) of 3.121. Note however, that the direction of the difference is reversed with respect to the result in engagement; for naturalness, it is loading spatial working memory which leads to lower scores.

Effect	**df**	**F**	***p***
Effective WM load	1	0.005	0.943
WM system	***1***	***4.332***	***0.038***

Table 5.6: GLM results for the ITC-SOPI naturalness factor (significant predictors highlighted). Intercept has been excluded for clarity.

5.8.3.4 ITC-SOPI Negative effects factor

The Cronbach's alpha for this factor was 0.89. For this factor, the main model is not significant: $F(2, 173) = 0.010$, $p < 0.989$. See Table 5.7 below for partial effect sizes (the intercept results have been excluded from the table for the sake of clarity).

Effect	**df**	**F**	***p***
Effective WM load	1	0.021	0.885
WM system loaded	1	0.004	0.949

Table 5.7: GLM results for the ITC-SOPI negative effects factor. Intercept has been excluded for clarity.

5.9 Discussion

5.9.1 Support for CLCC predictions

In terms of the predictions made by the CLCC model in 5.2 above, the results suggest the following:

1. *Performance on the main task will degrade as a function of loading task difficulty* – The negative relationship between task performance (number of doors opened inside the 15 minute period) and the effective working memory supports the prediction. This indicates that the experimental manipulation was successful (as measured by effective working memory load).

2. *Working memory load will negatively affect presence* – In general terms, the evidence does not support this prediction. Effective working memory load did not predict any of the four ITC-SOPI factors. The lack of a finding cannot be attributed to an ineffective manipulation of working memory load, because the effective working memory load measure successfully predicted task performance (see prediction 1 above). Furthermore, the extremely small effect sizes of effective working memory load on the four ITC-SOPI factors suggests that the lack of a significant finding is not simply an issue of low statistical power; it would seem that if the effect exists at all, it would be extremely small. This contradicts the CLCC prediction, which sees working memory allocation as playing a major role in presence. This would seem to suggest that if working memory is involved in presence, it either plays a minor role, or no role at all.

3. *Working memory load will affect cognitively higher forms of presence (such as engagement) more than spatial presence* – This prediction was partly supported. The spatial factor had a non-significant model, but the engagement and naturalness factors were predicted by the working memory system loaded. In both of these higher order forms of presence, loading one of the systems led to a modest increase (see 4 and 5.9.2 below). As with the previous prediction, the CLCC model suggests that a large difference in the amount of working memory used by these forms of presence should exist, but the data shows that the difference is quite small.

4. *Loading spatial working memory should impact spatial presence more than cognitively higher forms of presence (such as engagement), but loading verbal working memory should affect cognitively higher forms of presence more than spatial presence* – This prediction was partly supported by the data. Engagement showed a modest (but significant) interference effect when the verbal system was

loaded, as predicted by the CLCC model. Spatial presence showed no difference between the working memory systems loaded, while the higher order forms of presence (engagement and naturalness) showed a difference between the working memory systems loaded, but in the opposite direction predicted by the CLCC model.

5.9.2 Effect of working memory system loaded

Loading each of the working memory systems has an extremely interesting effect on the cognitively higher forms of presence (engagement and naturalness), which provides some support for the CLCC model. Engagement behaved as predicted by the model: due to the degree of semantic involvement (and therefore a need for phonological loop space as a processing space), rehearsal of the digit code produced an interference effect and lowered engagement scores. Unfortunately, there was no effect of effective load on engagement, which would have provided unequivocal support. Naturalness also produced an interesting interference effect, although for this factor it was loading *spatial* working memory that led to interference. This suggests that naturalness makes more use of spatial than of verbal working memory. This finding is in fact supported by the CLCC model – recall (see section 4.6 in chapter 4) that naturalness is thought to exist in the model as a set of implicit expectations in active memory clusters. If these expectations are met by incoming perceptual data, then a sense of naturalness in the VE will result. An examination of the model architecture reveals that the interface between perceptual data (arriving from the stimulus attenuator) and active knowledge clusters is the temporary structures in working memory. The data from this study suggests that the temporary structures which allow these implicit expectations to be met are primarily formed in spatial working memory.

5.9.3 Overall lack of working memory effect on presence

The lack of working memory effects on presence suggested by these data is surprising. Although working memory has not been incorporated into any previous presence models, this result does in fact contradict a large body of literature. Essentially, the design of this study introduces progressively more aggressive attention distractions. According to the environment selection model (Slater, 2002), distractions should increase the likelihood of a break in presence and lead to a reduction in presence. The three-pole model (Biocca, 2003) predicts that the mental rehearsal task will shift presence towards the mental pole, and thus away from a zone of undivided presence. According to Waterworth & Waterworth's FLS model (2001), a mental rehearsal task should lead to the focus dimension turning inwards, and due to the repetition of the rehearsal task, a reduction in sensus, which in turn will decrease focused presence. Finally, the MEC model (Wirth *et al.*, 2007) would predict that a distracter of this form would interfere with presence at two levels: first, it would pull away automatic attention and therefore interfere with SSM construction; and secondly, it would interfere with the attentive aspects of ERF hypothesis testing, and thus reduce the probability of the VE SSM being selected as the PERF. However, none of the four presence factors tested showed any differential effect as a function of working memory load. Although this may suggest that the study was somehow profoundly flawed, there is one study, by Lee, Kim & Lee (2004a) which partly corroborates the findings. In that study, subjects were given either given a counting task (similar to our phonological loop loading task) or not, and placed to navigate in one of three VEs, which differed in terms of fidelity. Subjects were measured using ten items from Witmer & Singer's Presence Questionnaire. Like out study, they found that sustained attention (manipulated by the counting task) showed no effect on PQ scores. Furthermore, Lee *et al*. suggest that although they only evaluate spatial presence, attention may have an effect on what they term "non-spatial" forms of presence (such as engagement) – which again is consistent with the data of this study. They argue that this is because spatial presence makes use of low-level neural circuitry, while "non-spatial" presence involves other brain regions (that is, low-level and high-level cognition). This is conceptually compatible with the CLCC model, if one accepts that spatial presence may be mostly a function of the folk physics modules (as argued by K. M. Lee, 2004), and other forms such as engagement and naturalness involve an interaction between declarative memory and perceptual data.

5.10 Conclusion

In general, the study does not find a great deal of support for the working memory aspects of the CLCC model. The study shows that working memory plays no role in spatial presence, and a small role in cognitively higher order forms of presence. This lack of effect cannot be attributed to an ineffective manipulation of working memory load, as task performance was affected by working memory load in the expected direction. However, the difference in difficulty between the spatial and verbal load conditions (the former loaded on average 0.850 of capacity while the latter loaded only 0.517 of capacity) is a problem. Presumably, the difference exists because of a sample effect. Because all the subjects used were psychology undergraduate students, it is likely that the amount of reading required by their courses has led them to develop strategies to maximize their verbal working memory (Baddeley, 2004; Carpenter & Just, 1989). Regardless of the reason, the difference between these conditions increases the difficulty of identifying working memory effects, because the average load of 0.517 may simply not be enough to introduce interference effects. Nevertheless, the phonological loop should play no role in spatial presence, so if working memory played a role in spatial presence, this should have appeared as an effective working memory load effect.

Chapter 6: Study 2 – Working memory use in a learned media decoder

The previous study examined the role of working memory in presence, and found a role for working memory in the presence experience: Engagement showed interference when the subject experienced a concurrent phonological loop load, and naturalness showed interference when the subject experienced a concurrent visuospatial sketchpad load. However, these findings were modest, and smaller than those predicted by the CLCC model. Furthermore, the CLCC model predicted that interference effects would scale with the size of the concurrent loading task, but this phenomenon failed to appear in those data, even though the experimental manipulation was effective.

One possible reason for the lack of a working memory effect in Study 1 was the type of media decoder being used by the subjects during the VE experience. Recall that the CLCC model proposes that all subjects have a set of media decoders which can be used to access the VE encoded in the medium. Some decoders are inherent and therefore highly resource efficient, requiring very little working memory to process the medium; while others are learned, and require more working memory for processing (see section 4.3.5 in chapter 4 for a full description of media decoders). In Study 1, the subjects used a real-time, interactive, high-fidelity graphical interface to interact with the VE. This situation was likely to engage an inherent visual media decoder (as the scene closely approximated decoding a normal retinal image during real navigation). Therefore, it is likely that the situation used in Study 1 led to subjects processing the scene with little use of working memory by virtue of the media decoder which the scene activated. This is a plausible explanation, consistent with the CLCC model, as to why so few working memory effects were evident in that study.

The present study aims to investigate if the lack of effect in Study 1 can be attributed to the media decoder which was active during the experience, or if it was due to a general lack of working memory effects on presence. In order to do this, this study replicates the method of Study 1 (in particular its use of a concurrent loading task), but attempts to stimulate a media decoder which is known to make extensive use of working memory during processing. This is done by presenting a VE in text form rather than in a high-fidelity graphical form; to navigate the environment, the subject must read text descriptions of the VE at their position. As reading is known to make extensive use of working memory (Baddeley, 1998; Carpenter & Just, 1989), presenting the subject with a text-based VE should create a situation similar to that used in Study 1, but with a higher working memory demands. This should, according to the CLCC model, accentuate any working memory effects on presence (see 5.2 below for a detailed description of this prediction).

Creating a text-based VE could give rise to some protest that presence can only arise from particular forms of media (see Slater, 2003a for a discussion of this objection) although the evidence suggests that presence can arise from text-based environments (Biocca, 2003; Nunez & Blake, 2003b; Towell & Towell, 1997). The CLCC model proposes that presence is a product of processing a VE which has been decoded from a medium, and therefore the form or medium in which the medium is presented is not an issue (evidence from Nunez & Blake, 2003b suggests that this is an empirically defensible position). Therefore, from the perspective of testing the CLCC model, the use of a text-based VE is justified.

6.1 Predictions about working memory in learned media decoders

5. *Performance on the main task will degrade as a function of loading task difficulty* – as performance on the main task (the door opening task used in Study 1, which involves rehearsal in working memory) is independent of the VE medium used, this study should reproduce the finding from Study 1.

6. *Working memory load will negatively affect presence* – The essential difference between this study and Study 1 is that subjects in this study will be making use of a learned decoder, which according to the CLCC model (see 4.3.5) will require more working memory than the inherent decoder used in Study 1. Therefore, one would expect that fewer resources would be available for creating temporary structures in working memory, which should result in a reduced presence experience. Theoretically, one would expect the reduction in presence scores in this study to be greater than those found in Study 1, but a direct comparison between the two studies is not possible due to extraneous (but unavoidable) differences between the two studies.

7. *Working memory load will affect cognitively higher forms of presence (such as engagement) more than spatial presence* – The CLCC model (see 4.3.5) predicts the only difference between learned and inherent media decoders is the amount of working memory they use during processing (at least for completely developed decoders; it is not clear from the model what differences exist when learned modules are in the process of being formed). Therefore, given that the working memory requirements of the different forms of presence are independent of the medium used to display the VE, one would expect the results of this study to follow the pattern of results found in Study 1.

8. *Loading the verbal WM system should impact all forms of presence more than loading the visual WM system* – Given that the media decoder required to process a text-based media decoder will make extensive use of the phonological loop during processing (Carpenter & Just, 1989), it is to be expected that a concurrent working memory load task which also uses this system will significantly affect the decoding process. As in Study 1, one would expect the forms of presence which make more use of semantic processing (engagement and naturalness) to be more severely affected than spatial presence; however, unlike Study 1, in this study the prediction must also be made that spatial presence will be reduced by the concurrent verbal loading task, as verbal working memory will be needed to decode the medium and construct a cognitive model of the space.

6.2 Sample

The study was made publicly available on the internet (see 6.3 and 6.4 for a description of the procedure used). The web-site associated with the study was advertised to several university humanities undergraduate classes by email and course forums, as well as by placing poster advertisements about the campus. A total of 114 subjects completed the study.

Because the data were not collected under controlled laboratory conditions, each subject's data was screened for possible errors (arising from software faults or other causes) before inclusion in the study. Two criteria were used for screening the data:

1. Only subjects with a measured working memory (using the Corsi block tapping test or the immediate recall digit span task – see 6.6. below) of less than 12 were included for analysis (this value is mentioned this as the extreme upper end of the normal range of working memory capacity by Baddeley, 1998; Vandierendonck *et al.*, 2004). This excluded 5 subjects (3% of the sample), whose measured working memory capacity scores ranged between 16 and 27, indicating a high probability of software failure or the use of some external mnemonic aid by the subject.

2. The VE exploration phase of the study was timed by the system and ended after 20 minutes. Due to software errors, some subjects experienced the VE for shorted periods of time. To control for possible effects of length of experience, we excluded all subjects who experienced the VE for less than 17 minutes – a total of 20 subjects.

After this screening, 89 subjects remained (38 men and 51 women). The mean age of the subjects was 20.6 years (S = 3.55). The mean working memory span was 6.27 chunks (S = 2.25).

6.3 Apparatus

This study was opened to large-scale voluntary participation over the internet. It was advertised as a study on the role of memory in learning new places. The text-based virtual environments used (see 6.3.1 below), instruction slides and electronic questionnaires were packaged into an automatic installer which was made available on the World Wide Web (the package can be downloaded from http://chomsky.uct.ac.za/dnunez/phd/study2.zip). The study was advertised using posters on a university campus, and subjects were encouraged to pass the information on to acquaintances.

When the study package was downloaded, it installed and ran the study automatically. As the environments were displayed with text, there was not much concern about standardizing physical display size; also, there was no requirement of specialized graphics or sound rendering hardware. The application was designed to run on a basic installation of Windows XP, on an 800 x 600 x 32 full-screen mode, using anti-aliased fonts (see 6.3.1 below for screenshots). Once the application had run the entire study (see 6.4 below, for the procedure), it uploaded the data automatically to a server, and shut itself down. The application left an entry in the Windows registry to prevent itself from being run more than once on the same computer, as a control against a single subject providing more than one data point.

6.3.1 Virtual environments

This study used a text-based VE system similar to that used by Nunez & Blake (2003b). The VE is divided into rooms, with a number of exits (which occur as North, South, East, West, upstairs or downstairs in the interface) connecting the rooms. Each room is described by a short piece of text, which contains spatial information, as well as descriptions of any sounds present. The descriptions often include impressions of the room (e.g. 'this wing of the building feels cold and deserted'). Interaction is by mouse clicks on a menu bar at the bottom of the screen (see Figure 6.1 and Figure 6.2 below). In any room, subjects can move to another connected room, or access the panel and locks associated with the working memory task (see 5.6 below). These actions are also the only basic actions available in the VE system used in Study 1 (see 5.4.1 in chapter 5), thus maintaining as much similarity as possible between the two systems.

Two environments were used in the study – a training VE and the main VE. The training environment (shown in Figure 6.1 below) gave a tutorial of how to navigate through the VE, and how to complete the working memory task (described in 5.6 below). The training VE ended when the subject successfully navigated all the rooms in the VE, including making use of the panels and locks from the working memory task. The main VE implemented the floorplan of the hospital VE used in Study 1, with the locks and panels associated with the working memory task placed in the same approximate positions as they were in Study 1, again to maintain similarity with that study. The main VE is shown in Figure 6.2 below.

6.4 Procedure

The subjects entered the study by visiting the web-site. On the site, the subjects were given a brief introduction to the study (that it was a study on the role of memory in visiting a new place), and indicated consent for participating in the study by clicking a link which began the downloading of the study package. Once the package was downloaded, it was automatically installed and run. The subjects then viewed a series of slides explaining the task and interface of the text-based VE, and were taken into the training VE. When the subject completed the training VE, they were given the same scenario for the upcoming task as used in Study 1 (see section 5.6 in chapter 5, and appendix E for the scenario text).

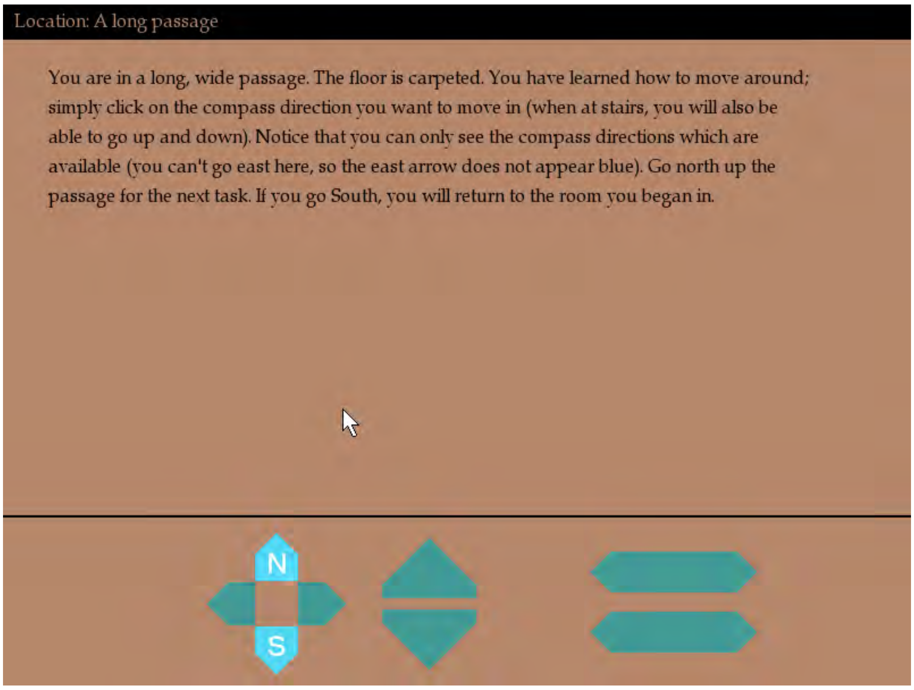


Figure 6.1: The text-based training environment. The top bar shows the title of the room, while the main area of the screen shows the description. The icons in the bottom area are for interaction (notice only possible exits, N and S in this case, are shown, while the others are ghosted out).

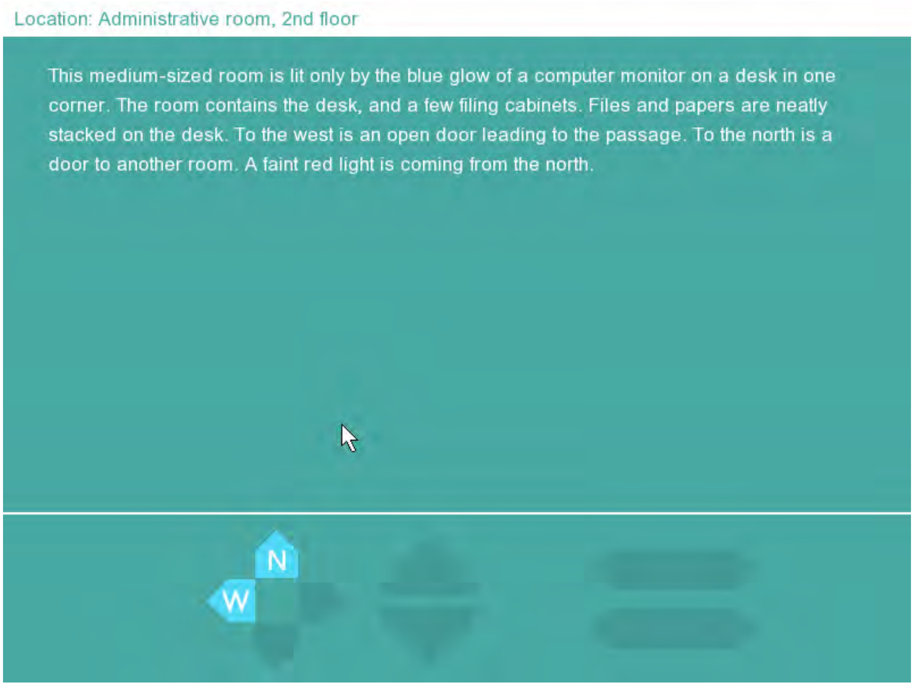


Figure 6.2: The main hospital VE in a room with no lock and no panel. Possible exits are N and W.

6.5 Experimental task

The task used in this study is the same used in Study 1 (see 5.6 in chapter 5). The subjects were given a working memory loading task (rehearsal and recall of lists of random strings of information) to perform during the VE experience, based on the tasks used by Baddeley *et al.* (Baddeley, 1986; Baddeley *et al.*, 2002). Different task conditions controlled the amount of working memory available to process the VE during exposure. Two versions of the task were created, one for each of the two working memory systems – one loaded the spatial system (the visuospatial sketchpad), and the other loaded the verbal system (the phonological loop). Each of these tasks was crossed with three different levels of load (3, 4 and 5 chunks), to create a 2x3 factorial design. Note that this differs from Study 1, where 4 levels of working memory load were used (see Table 5.3 below for the number of subjects used in each condition). A goodness-of-fit analysis (see 6.7.1 below) shows that the subjects were evenly distributed in the design.

The decision to use one fewer level of working memory load was based on the desire to reduce the required sample size for the study (by eliminating two cells), and the observation in Study 1 that no difference existed between the 5 and 6 levels of working memory load. Both factors (WM load and WM system loaded) are between-subject factors, with each subject contributing data to only one of the six cells.

WM system loaded	***WM load (chunks required to complete task)***		
	3	***4***	***5***
Verbal (phonological loop)	7	10	20
Spatial (visuospatial sketchpad)	15	17	20

Table 6.1: Study design and number of subjects randomly assigned into each conditions

As in Study 1, the task was to navigate through the environment by opening locked doors. The VE was created so that only one particular path led from the starting point on the top floor to the exit in the basement level. Along this path, 16 locked doors were placed. As with Study 1, the subject had to search for the panels; obtain the code, and rehearse it in working memory while they navigated to the door, and then use the code to open the door. If the subject had forgotten the code when they reached the door, they had to return to the panel to get the code again. The VE was structured so that there was only ever one available panel for the next locked door – this way, subjects only had to rehears a single code at a time. Panel access and code input was done by clicking on on-screen buttons with the mouse. When in a room with a panel or lock, the appropriate button became un-ghosted on the interface, and subjects could view or enter the code as in Study 1 (see Figure 6.3 and Figure 6.4 below). In the case of the verbal condition, a different font was used in the panel and the lock; this ensured that subjects could not used the shape of the numbers (and thus make use of spatial working memory) to recall the code. No digit was used more than once in a code. Figure 6.3 below shows the verbal condition panel interface, and Figure 6.4 shows the verbal condition door interface. The VE system recorded the number of successful and failed attempts at opening each door as a measure of task performance.

Figure 6.3: A panel showing a code of length 5 for the verbal load condition (notice that this system uses the same graphics and fonts used in Study 1) .

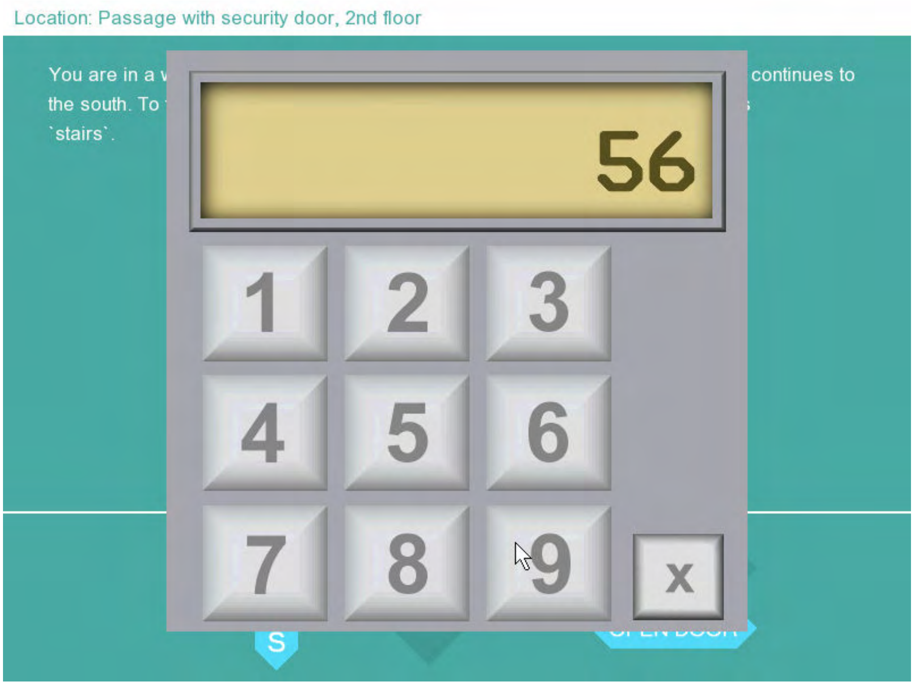


Figure 6.4: The lock in the verbal load conditions, with a code being entered. Note a different font is used than in the panel, to prevent visual rehearsal (note the same interface is used as in Study 1).

In the spatial load condition, a square grid of 3x3 buttons was used (see figures 5.6 and 5.7 in chapter 5). In the panel, the code sequence was played back (each button was illuminated for one second, with a 250 millisecond delay between buttons), and the subject could watch the sequence as many times as they wished by clicking the 'show' button on the panel. No block was used more than once in the code sequence. In the door interface, the subject had to repeat the sequence in order to open the door – buttons remained lit as the subject clicked the code. The interface for the spatial load condition was the same as used in Study 1.

6.6 Measures

The main dependant variable (presence) was measured using the ITC Sense of Presence Inventory (ITC-SOPI - Lessiter *et al.*, 2001). The choice of this measure is discussed in 2.6 in chapter 2.

Due to individual differences in working memory capacity across individuals (Cowan, 2001), subjects had their working memory capacity evaluated with either the digit span task with immediate recall (Waters & Caplan, 2003) or the Corsi block-tapping test with immediate recall (Vandierendonck *et al.*, 2004). These measures are discussed in section 5.7 in chapter 5. Unlike Study 1, where subjects had both spatial and verbal working memory measured, in this study, subjects were only measured for the working memory capacity relevant to the condition they were placed in (e.g. a subject in the verbal load condition would only be given the digit span measure, and not the Corsi block tapping measure). This was done for two reasons: first, due to the lack of predictive power of cross-modality working memory measures found in Study 1, and second, in the interests of keeping the study as short as possible for the subjects. As with Study 1, all measures were implemented as computer-based tests, and automatically administered immediately at the end of the hospital VE experience.

6.7 Analysis & results

6.7.1 Allocation of subjects to design

The system randomly allocated each subject to one of the study's cells, which resulted in an even distribution across the conditions (Table 5.3 above; $\chi^2 = 2.26$, df = 2, $p < 0.323$). Although women outnumber the men in the sample as whole (precluding the examination of gender effects), the men were evenly distributed among the cells of the design ($\chi^2 = 0.365$, df = 5, $p < 0.832$).

6.7.2 Evaluation of experimental manipulation

As in Study 1, we wanted to establish that the working memory manipulation was successful before further analysis. As Study 1 showed that the effective working memory load (i.e. load imposed by the task divided by the subject's measured working memory span) was a better predictor than the raw task difficulty, only effective working memory load was considered. A linear regression to predict task performance (number of doors opened during the 20 minute VE experience) from effective working memory load shows a small but significant relationship ($F(1,87) = 75.157$, $p < 0.019$, $R^2 = 0.06$) – see Figure 6.5 below for the scatterplot. Note that the beta is negative (-0.25), indicating that as the effective working memory load increased, task performance decreased. This indicates that task performance is being interfered with by the concurrent working memory loading task, as predicted by Baddeley (1986). This finding concurs with that in Study 1, and shows that controlling for individual differences in working memory capacity, the locks and panels task is an effective working memory interference task.

6.7.3 Effect of working memory load on presence

As in Study 1, the data were analyzed with a set of general linear models (Hastie & Tibshirani, 1997; Neter *et al.*, 1988). Each of the ITC-SOPI factors was modeled using the working memory system being loaded (spatial or verbal), and the effective working memory load (see 5.8.3 in chapter 5 for a justification of this analysis technique).

6.7.3.1 ITC-SOPI spatial factor

The Cronbach's alpha for this factor was 0.92. For this factor, the overall model is not significant: $F(2, 86) = 0.455$, $p < 0.635$. See Table 5.4 below for partial effect sizes (the intercept results have been excluded from the table for the sake of clarity). This finding concurs with that found in Study 1.

Effect	df	F	*p*
Effective WM load	1	0.908	0.343
WM system	1	0.286	0.593

Table 6.2: GLM results for the ITC-SOPI spatial factor. Intercept has been excluded for clarity.

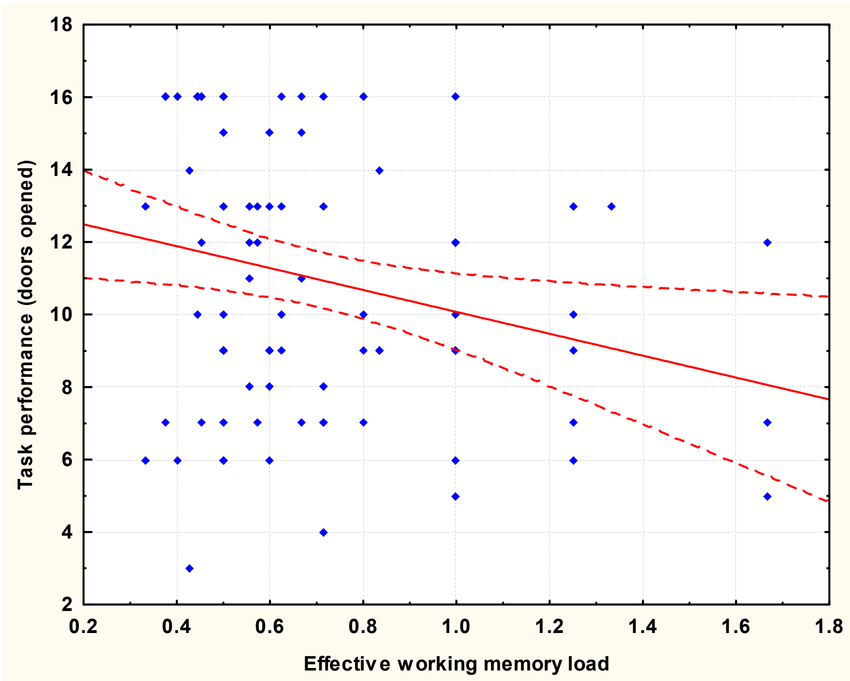


Figure 6.5: Scatterplot of effective working memory load versus task performance. Note that as effective load increases, task performance decreases, indicating an interference effect.

6.7.3.2 ITC-SOPI engagement factor

The Cronbach's alpha for this factor was 0.83. For this factor, the overall model is not significant: $F(2, 86) = 1.398$, $p < 0.252$. See Table 5.5 below for partial effects (the intercept has been excluded from the table for the sake of clarity). This finding differs from that in Study 1, which showed that WM system loaded was a predictor of engagement.

6.7.3.3 ITC-SOPI naturalness factor

The Cronbach's alpha for this factor was 0.77. For this factor the overall model is again not significant, showing an unusually small effect: $F(2, 86) = 0.006$, $p < 0.993$. See Table 5.6 below for partial effects (the intercept results have been excluded from the table for the sake of clarity).

Effect	df	F	*p*
Effective WM load	1	2.412	0.124
WM system	1	0.073	0.787

Table 6.3: GLM results for the ITC-SOPI engagement factor. Intercept has been excluded for clarity.

This finding differs from that of Study 1, which showed that WM system loaded was a significant predictor. It should be noted that the partial effect of WM system loaded in this model is significant, but as the overall model is not significant, this finding cannot be taken as definitive evidence supporting the findings in Study 1.

Effect	df	F	*P*
Effective WM load	1	0.005	0.943
WM system	1	4.332	0.038

Table 6.4: GLM results for the ITC-SOPI naturalness factor. Intercept has been excluded for clarity. Note that although the partial WM system effect is significant, this should not be given any evidentiary weight, as the overall model is not significant.

6.7.3.4 ITC-SOPI Negative effects factor

The Cronbach's alpha for this factor was 0.82. For this factor, the main model is not significant: $F(2, 86) = 2.587$, $p < 0.081$. See Table 5.7 below for partial effect sizes (the intercept results have been excluded from the table for the sake of clarity). This finding concurs with Study 1.

Effect	df	F	*P*
Effective WM load	1	1.689	0.197
WM system loaded	1	0.862	0.355

Table 6.5: GLM results for the ITC-SOPI negative effects factor. Intercept has been excluded for clarity.

6.8 Discussion

6.8.1 Support for CLCC predictions

In terms of the predictions made by the CLCC model in 5.2 above, the results suggest the following:

1. *Performance on the main task will degrade as a function of loading task difficulty* – This prediction was supported, as evidenced by the negative slope of the regression (see 6.7.2 above). This shows that the experimental manipulation was successful.

2. *Working memory load will negatively affect presence* – As with Study 1, there is no evidence to suggest that working memory load affected presence, as none of the four ITC-SOPI factors were predicted by the effective working memory load. As with Study 1, this lack of effect cannot be attributed to an unsuccessful manipulation of the working memory load, as the load manipulation was indeed successful (see prediction 1 above). In Study 1, it was hypothesized post-hoc that the lack of working memory effect may be due to the fact that an inherent media decoder uses almost no working memory; however, in the case of this study, it is already established that decoding text requires working memory and is subject to interference effects (Baddeley, 1998; Carpenter & Just, 1989). Therefore, if presence shows no working memory interference when using the text decoder, then one must be inclined to conclude that presence does not make use of working memory, regardless of whether the medium requires decoding in working memory or not.

3. *Working memory load will affect cognitively higher forms of presence (such as engagement) more than spatial presence* – This prediction is related to prediction 2 above. Unlike Study 1, where there was a small effect of system loaded on engagement and naturalness, in this study, there was no difference between the cognitively higher (engagement and naturalness) and lower (spatial and negative effects) forms of presence. The prediction is therefore not supported. A small caveat exists however. The GLM for naturalness, while not significant, showed

the slightly incongruous finding that WM system loaded was a significant predictor (in the same direction as in Study 1). This may indicate that a very small WM system effect exists for naturalness, although it may also be a spurious finding.

4. *Loading the verbal WM system should impact all forms of presence more than loading the visual WM system* – This prediction was not supported. There was no difference on any of the four ITC-SOPI factors in terms of WM system loaded. This again contradicts the findings of Study 1, where small but significant differences were found between loaded systems on presence.

6.8.2 Lack of working memory effect on presence

In general, this study supports the lack of working memory effect on presence found in Study 1; however, while Study 1 showed small differences (for the system loaded), this study shows almost no significant difference for effect loaded (with the exception of the indication of a possible effect for naturalness), and no effects whatsoever for effective working memory load. This finding is very surprising in terms of the CLCC model. According to the model, inherent media decoders (such as the vision decoder tested in Study 1) should use very little working memory, thus making the findings of Study 1 not entirely surprising. However, the model clearly predicts that learned decoders which require significant amounts of working memory (such as used in this study) should show *more* severe presence interference effects, not less, as found in this study.

This counter-intuitive finding cannot be explained as occurring due to an error in experimental manipulation for two reasons. First, the regression analysis of task performance versus effective working memory load shows the expected negative slope, indicating that the task was indeed harder in the higher load conditions. Second, even if the task difficulty manipulation had not been successful, the manipulation of which working memory system was loaded must have shown some effect, as this was essentially a replication of Baddeley's classic method used to establish the existence of two separate working memory systems (Baddeley, 1986, , 1998).

Nevertheless, before one concludes that the CLCC model is fatally flawed, there are three possible flaws in the design of the study which may have masked a working memory effect. The first flaw is a lack of control for the efficiency of the text decoder, as no measure was taken of the reading skill of the sample. Ideally, one would have established how much working memory should have been loaded for each subject or how much working memory could be loaded before reading exhibited interference effects, and then run experiment starting at this level of load. However, this measure would only have reduced the degree of error variance in the linear models, as subjects were randomly assigned to conditions, and therefore no systematic differences could have been introduced into the experiment this way.

A second flaw, related to the first, is that the study only used modest levels of load (3 to 5 chunks), and interference effects might only appear at higher levels of effective load. Nonetheless, as Figure 6.5 above shows, most subjects were operating at the 0.6 to 0.8 level of effective load, which suggests that this was indeed a high level of load for this sample. Also, even if the levels of load were modest, study 1 showed differences between the systems loaded at the same levels of load used in this study, suggesting that some (weak) effects are possible at these levels of load.

The third and final flaw in the study is the modest sample size. Study 1 showed, with a slightly larger sample, weak effects on presence. As it was expected that the effects with the learned decoder would be larger, it was deemed acceptable to use a smaller sample; however, it seems that the effects, if they exist at all, are very small so a larger sample may have been able to identify them.

6.8.3 Differences between inherent and learned media decoders on presence

The purpose of this study was to examine possible differences in presence when the VE is decoded by an inherent media decoder, as opposed to a learned media decoder. Study 1, which examined the effects of working memory on presence when using an inherent decoder found that working memory had a small influence on presence. In particular, cognitively higher forms of presence (engagement and naturalness) showed a system effect, with engagement showing verbal interference and naturalness showing spatial interference. Most significantly, there was no effect of working memory load on presence, indicating that contrary to CLCC model predictions, presence does not require a large amount of working memory for temporary constructions. The present study, which examined the same predictions as study 1 but in the context of a learned media decoder, showed no working memory effects at all; the small effects found in study 1 were not replicated.

There is however one important similarity between the two studies, which may inform the future development of the CLCC model: No working memory load effects were found in either study. Although one cannot lightly infer that a lack of a significant effect in an experiment means the effect in question does not exist (Rosenthal & Rosnow, 1991; Turner & Roth, 2003), in this case one may be able to infer from the similarities between the two studies that there is indeed no working memory effect in presence for the following three reasons. First, both studies find extremely small effect sizes with moderately sized samples. This suggests that independent of possible study artifacts, *the* effects are extremely small. Second, both studies use different samples, different media, and different interfaces to the VE; the lack of effect therefore cannot be attributed to the artifacts of any one study, suggesting rather than presence is not subject to working memory effects. Finally, the two studies use a small variation of the method used by Baddeley, which showed large interference effects for a range of other cognitive tasks. The lack of an effect for presence with the same method suggests that no effect exists to be found.

6.9 Conclusion

This study confirms the general finding of Study 1 that working memory plays very little role in presence, and suggests that there may be no working memory effects on presence at all. As with Study 1, the lack of effect cannot be attributed to a weak experimental manipulation. However, unlike study 1 which found some small effects for the cognitively higher forms of presence, this study finds no interference effects at all. Taking the findings of both studies together, one must conclude that working memory plays a far smaller role in presence than predicted by the CLCC model.

While this is not a surprising finding in Study 1 (as that experiment likely stimulated the use of an inherent media decoder), it is far more surprising in the context of this study, which stimulated the use of a learned media decoder. The implications for the

CLCC model are clear: if constructions do happen in the model, then they are not held in working memory. Furthermore, because there was no difference in effect when using an inherent or learned media decoder, it would seem that decoding the VE and constructing the experience do not share a single limited pool of cognitive resources as suggested by the model. It is likely that the decoding happens in working memory (as this task shows interference effects), while construction associated with the presence experience occurs further upstream (perhaps with declarative memory coming into play at an earlier stage in the process than suggested by the CLCC model). Most interestingly, these two experiments suggest that distraction effects (as discussed in Freeman *et al.*, 2000; Slater & Steed, 2000; Wirth *et al.*, 2007) do not occur as a consequence of interference effects as suggested by the CLCC mode (see section 4.4.2 in chapter 4); rather, they must predominantly occur as a function of the stimulus attenuator (see 4.3.2 in chapter 4).

Chapter 7: Study 3 – Relative contribution of working memory and attention to presence

The previous two studies have examined the role of working memory in presence, and generally found there to be only small effects, both on an evolved, inherent decoders (such as vision), or on learned decoders (such as reading). Although these studies covered some of the major functions associated with the bottom-up components of the CLCC model (which process data between the environment and working memory), they contained one important theoretical weakness. Because the stimulus attenuator exists upstream of working memory in the only path between the environment and working memory, both previous studies conflated possible effects of the stimulus attenuator and working memory. The present study aims to separate out the effects of working memory and attention (conceptualized in the CLCC model as the stimulus attenuator), to examine if they have differential roles in presence, and therefore if the separation between these components which exists in the CLCC model is empirically justified.

7.1 A dual-task method for comparison

Previous studies which have attempted to evaluate the contribution of working memory to a task have given the subject two tasks: the task of interest, and a second working memory loading task. If the second task produces an interference effect, then the two tasks must use the same pool of resources (see Study 1 in chapter 5 for a detailed description of this paradigm). A similar dual-task method exists in attention research, where subjects are given two simultaneous tasks which require them to divide their attention (for example the classic dichotic listening task of Broadbent, 1954; or the visual analog used by Treisman, 1964). As with working memory loading studies, the logic of the design is that if the two tasks make use of the same pool of attentional resources, then there will be an interference effect evident.

In order to decide if working memory or attention have the greater contribution to presence, the current study makes use of two dual task conditions (a working memory loading condition, and an attention loading condition), which are compared to a baseline, single-task condition. In the baseline condition, subjects view a first-person perspective video of a walk through a virtual environment with a simple visual task (pressing a key whenever the video shows walking through a door), and are then evaluated for their presence. In the attention loading condition, subjects are shown the same video with the same task, plus a second visual task which requires them to respond (with a different key press) when a blue square appears on a rapidly changing mosaic displayed adjacent to the video. Finally, in the working memory condition, subjects are shown two first-person perspective videos of walking through two VEs, and must press different keys when each video shows walking through a door. In the attention condition, subjects need only divide their visual attention to complete both tasks, and hence it loads only the attention system. In the working memory condition however, subjects must form complete and coherent mental models of both VEs to complete the tasks, and also divide their attention between them. Hence, this condition loads both the attention and working memory systems. By comparing the results against the baseline, it is possible to determine the relative contributions of each system to presence, thus:

1. If all three conditions produce equal presence scores, then neither attention nor working memory contribute to presence

2. If the working memory and attention conditions are equal but lower than the baseline, then attention contributes to presence, but working memory does not (recall that the working memory condition actually loads both attention and working memory)

3. If a stepped profile occurs (baseline gives higher scores than attention which in turn gives higher scores than working memory), then attention contributes to presence, and working memory contributes independently of attention.

Although other combinations are possible (e.g. attention being higher than baseline, or working memory being higher than attention), these are not theoretically possible, and so will not be considered for the time being (if they do occur, it will likely indicate an error in the design, and not a real effect).

7.2 Predictions about working memory, attention and presence made by the CLCC model

This study makes the following two predictions about the relative contribution of working memory and attention to presence:

1. *A concurrent task (be it a working memory or and attention loading task) will lead to a reduced presence experience* - although the outcome for a concurrent working memory and attention loading task is predicted to be the same, the reasons for the reduction in presence are slightly different. In the case of attention, the reduction in presence comes from a reduction in the VE relevant stimuli which are allowed in for processing. This is because the subject in the attention loading condition must divide attention and allocate some attention resources to the loading task. This will lead to less bottom-up information being available in working memory, and thus a reduction in the number of temporary structures for processing the VE. In the case of working memory, the divided attention problem remains (the subject must still divert attention between two tasks), but the subject also needs to construct two world models in working memory, which will further reduce the amount of temporary working memory structures available for the processing of the VE of interest.

2. *A working memory loading task will reduce presence more than an attention loading task* – Recall that the CLCC model proposes that presence depends largely on the amount of working memory allocated to creating temporary working memory structures, as these allow the connection between external stimuli and semantic meaning. Under an attention loading task, all working memory capacity is available for processing the environment of interest, but the amount of external stimuli related to the environment which makes it through to working memory is reduced. Conversely, under a loaded working memory system, a dual interference effect exists. First, the subject must split their attention between two tasks, so the working memory load task imposes a similar load to an attention loading task. Second, the subject needs to construct two environments in working memory to complete both tasks. This means that not only is there reduced bottom-up data with which to construct temporary structures, but there is also a

reduction in the amount of working memory available for the construction of temporary structures in working memory due to the competing requirements of processing the second environment. One would therefore predict that the presence experience for subjects in the working memory condition would be reduced relative to those subjects in the attention loading condition.

7.3 Sample

The experiment was offered to undergraduate psychology students as an elective for course credit. A total of 46 students participated - 34 women and 12 men. The mean age of the subjects was 21.69 years (S = 3.72).

7.4 Apparatus

The study was run on a single desktop computer in a cubicle which could be separated from the rest of the laboratory by thick ceiling-to-floor length curtains. The desktop computer used played the videos used in the study at a constant update rate of 30Hz, and displayed them on a 17 inch TFT LCD display (8ms grey-to-grey latency) at its native resolution of 1024 x 768 x 32. Subjects used the keyboard for input. A single subject was run at a time.

7.4.1 Virtual environments

This study used two videos of walkthroughs through two VEs. The VEs used for this purpose were the hospital VE used in Study 1, and the monastery VE used in Study 6, with minor changes. In the hospital VE, the doors and panels associated with the working memory task were removed, and in the monastery, the books associated with the object search and collection task were removed. The videos had a resolution of 640 x 480, no sound, and were encoded into DivX-4 format, using a 768 kbp/s encoding rate, which gave smooth, high quality playback at 25Hz. Each walkthrough lasted for 7 minutes. The walkthroughs represented exploration of the VEs (with pausing to examine objects, etc) rather than simply walking non-stop through the environments. The frequent stops and pauses in the path made it more difficult to predict motion than a simple end-to-end walkthrough.

To train the subject in the task, a set of slides were prepared (see Figure 7.1 and Figure 7.2 below for examples) and shown to the subjects prior to the experiment commencing. Subjects were also encouraged to ask the experimenter questions before the study began.

Figure 7.1: Sample of one of the instruction slides (for the working memory condition). The instructions showed still images, with no animation.

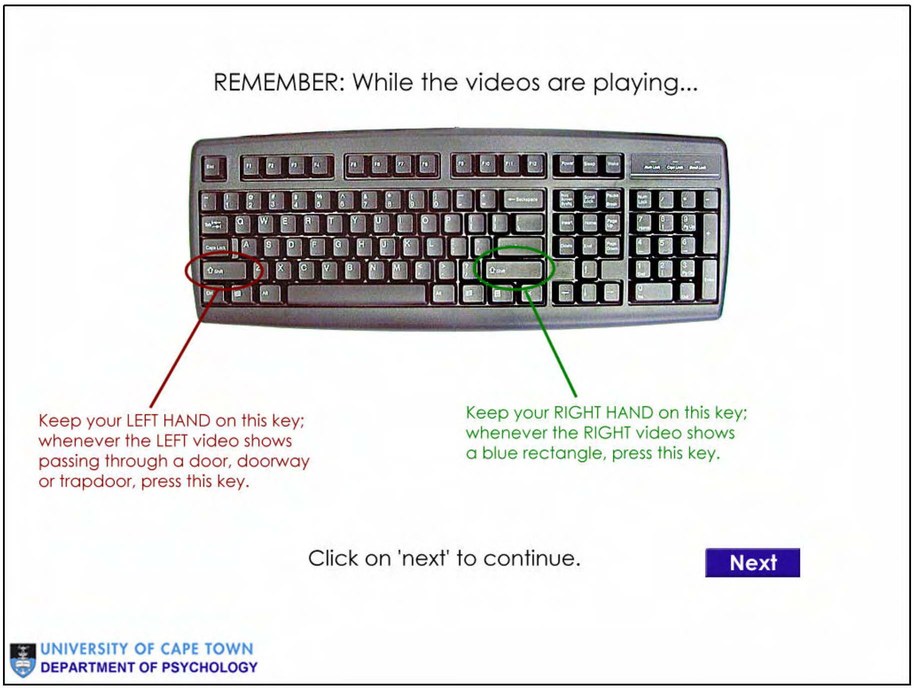


Figure 7.2: Sample of instructions (for the attention condition). This final reminder of the input was given to subjects just before the task began.

7.5 Procedure

The experiment ran over a period of three weeks. Subjects arrived at the study venue, and were greeted by the researcher. The researcher explained that the study was looking at psychological aspects of media processing. Subjects were then randomly assigned to one of the three experimental conditions (see experimental task in 5.6 above for an explanation of the conditions). Subjects were seated in front of the computer, and shown the instruction slideshow. The instructions differed for each of the three conditions, and are described in section 5.6 below.

When the slideshow was completed, subjects were asked if they had any questions, and these were responded to. The videos were then run. After a timed period of seven minutes, the videos terminated automatically, and the ITC-SOPI was administered on computer (see measures in 6.6 below). Subjects were instructed to respond to the ITC-SOPI only in relation to the hospital VE (which was displayed on the left of the screen in all conditions). The subjects were then thanked and the experiment was concluded.

7.6 Experimental task

The task for subjects was to press a key on the keyboard in reaction to particular events on the videos. All conditions lasted for seven minutes. There were three variations on the task:

Baseline condition: Subjects were shown a single video of a walkthrough the hospital VE taking up 25% of the screen area on the left of the screen (see Figure 7.3). The task for the subjects was to focus on the video, and press the left shift key on the keyboard whenever the video showed passing through a doorway.

Working memory condition: Subjects were shown two videos simultaneously, each taking up 25% of the screen area. The left video showed a walkthrough of the hospital VE, while the right video showed a walkthrough of the monastery VE (see Figure 7.4). The virtual walking speed on both videos was the same. The subjects were told to split their attention between the two videos, and perform two tasks simultaneously: When the left video showed passing through a doorway, they must press the left shift key; and when the right video showed passing through a doorway, they should press the right shift key.

Figure 7.3: Screenshot of the baseline condition (one video only).

Figure 7.4: Screenshot of the working memory condition.

Attention condition: Subjects were shown two videos simultaneously, each taking up 25% of the screen area. The left video showed a walkthrough of the hospital VE, while the right video showed a walkthrough of the monastery VE which was made incoherent by dividing it into 25 blocks which were randomly shuffled (see Figure 7.5).

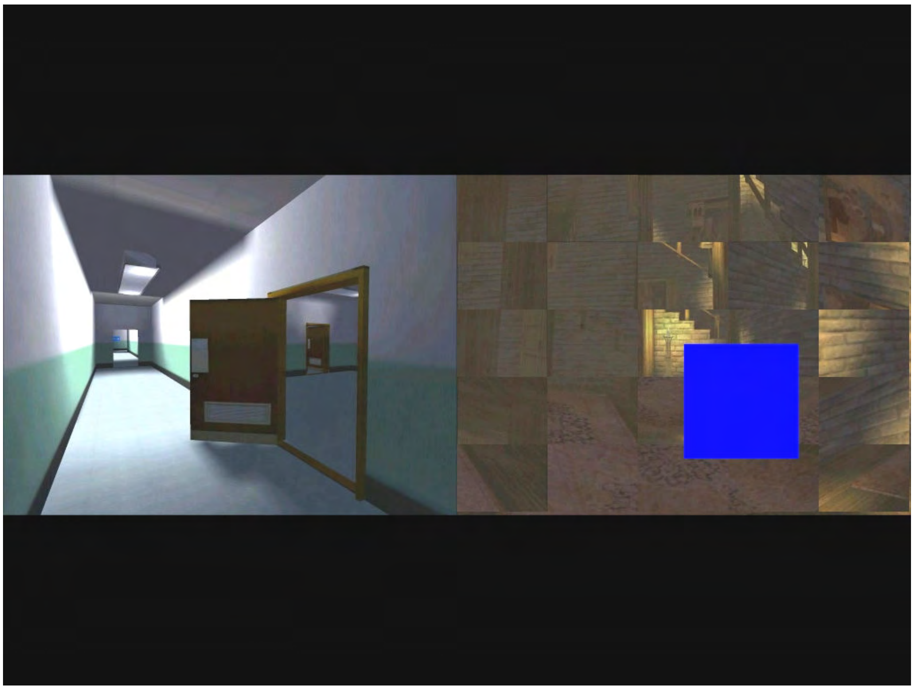


Figure 7.5: Screenshot of the attention condition (the right video was made incoherent). The blue square is the cue for the subject to press the right shift key.

This provided the same visual information (in terms of number and colour of pixels) as the working memory condition, but the shuffling made it impossible to infer a mental model from the video. The subjects were told to split their attention between the two videos, and perform two tasks simultaneously: When the left video showed passing through a doorway, they must press the left shift key; and when a blue square appeared somewhere on the video (see Figure 7.5), they should press the right shift key. The blue squares were timed to appear at the same time as the monastery video showed walking through a doorway (thus maintaining equivalence with the working memory condition), and remained on screen for 750 milliseconds.

7.7 Measures

The main dependant variable (presence) was measured using the ITC Sense of Presence Inventory (ITC-SOPI - Lessiter *et al.*, 2001). The choice of this measure is discussed in 2.6 in chapter 2. It is important to note that subjects in the working memory and attention conditions were instructed to respond to the ITC-SOPI only with respect to the hospital VE, and not to the monastery. They were only told this when they were given the questionnaire, after viewing the videos.

Performance on the task was measured by the number of key presses made by the subjects in response to events (passing through a doorway or seeing the blue block). A key press was only counted if was made while the subject was in fact inside the doorway. Two such measures were used: The first corresponds to the number of left key presses (in response to events in the hospital VE), which was measured for all three conditions. The second measure was number of right key presses (in response to events in the monastery VE), which was only taken for the working memory and attention condition (the baseline condition only showed one video, so no right key pressing task was assigned). Because the tasks involved only watching and no interaction, it was not deemed necessary to gauge the level of VR or video game experience of the sample.

7.8 Analysis & results

7.8.1 Allocation of subjects to design

The subjects were randomly assigned to the three conditions of the study, which resulted in an even distribution across the conditions (see Table 7.1 below). Although the sample contained more women than men (73.9% of the sample were women), the men were evenly distributed across the three conditions ($\chi^2 = 0.706$, $p < 0.702$).

Baseline	Working Memory	Attention
15	16	15

Table 7.1: Allocation of subjects to conditions (cell sizes)

7.8.2 Task performance

Both measures of task performance (left and right key press measures) were analyzed to look for manipulation effects using a one-way analysis of variance (ANOVA).

7.8.2.1 Left key presses

The ANOVA shows no significant differences between the three conditions: $F(2, 43)=1.1113$, $p < 0.338$. Figure 7.6 shows the means profile for this analysis.

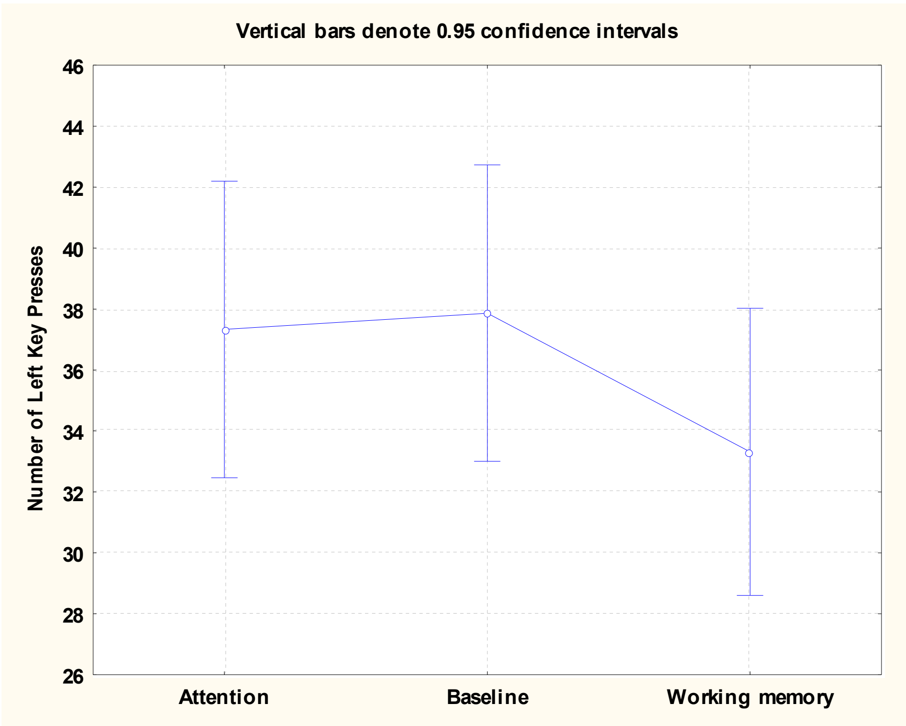


Figure 7.6: Average number of left key presses by condition

7.8.2.2 Right key presses

Only the working memory and attention conditions could be compared in terms of right key presses, as the baseline condition had no right key pressing task. The ANOVA shows a significant result, with the attention condition (M=35.733) resulting in more key presses than the working memory condition (M=31.875): $F(1, 29) = 8.5115$, $p < 0.007$.

7.8.3 Manipulation effects on presence

The four ITC-SOPI factors were tested for differences among the three experimental conditions using a one-way ANOVA. Recall that the IC-SOPI scores refer only to the left (hospital) VE. No significant differences existed for any of the four ITC-SOPI factors (see Table 7.2 below).

ITC-SOPI factor	df	F	*p*
Spatial	2, 43	0.163	0.983
Engagement	2, 43	0.056	0.944
Naturalness	2, 43	0.218	0.804
Negative effects	2, 43	0.219	0.804

Table 7.2: ANOVA results for ITC-SOPI factors across experimental conditions.

7.8.4 Presence and task performance

Although this study holds no predictions about presence and task performance, the performance of visual identification tasks is known to show interference effects with both working memory and attention loading (Baddeley, 1986; Treisman, 1969). Task performance may therefore be used as an indirect measure of attention or working memory contribution to a task. It was decided to use the common assumption that presence is a predictor of task performance proposed by, among others, Bystrom *et al* (1999a) and Youngblut and Hurie (2003). The model uses the four ITC-SOPI factors and the experimental condition as predictors of task performance using a general linear model (GLM). Only the left key press task was analyzed in this way, because analysis of the right key press requires the exclusion of the baseline task, which would a leave only 31 subjects, an unacceptably small sample for a GLM analysis (Hastie & Tibshirani, 1997; Neter *et al.*, 1988).

The overall model for left key presses is significant: $F(6, 39) = 2.425$, $p < 0.043$, $R^2 = 0.571$. An analysis of the partial effects shows that only the spatial and naturalness ITC-SOPI factors are predictors of left key presses. Table 7.3 below shows the effects (the intercept has been excluded form the table for the sake of clarity).

Effect	df	F	*p*
Spatial	***1***	***4.352***	***0.043***
Engagement	1	1.741	0.194
Naturalness	***1***	***9.607***	***0.003***
Experimental condition	2	1.326	0.277

Table 7.3: GLM results for left key presses. Significant effects ($p < 0.05$) are highlighted

An analysis of the partial correlations shows that spatial has a positive relationship to task performance (partial r = 0.318, $R^2 = 0.667$), while naturalness has a negative relationship to task performance (partial r = -0.447, $R^2 = 0.588$).

7.8.4.1 ITC-SOPI reliability

Cronbach's alpha coefficients were computed for each of the four ITC-SOPI factors. For Spatial, alpha was 0.91; for Engagement 0.83; for Naturalness 0.69, and for Negative Effects 0.85.

7.9 Discussion

7.9.1 Support for CLCC predictions

In terms of the predictions made by the CLCC model in 5.2 above, the results suggest the following:

1. *A concurrent task (be it a working memory or and attention loading task) will lead to a reduced presence experience* – No significant difference was found on left key presses between the three load conditions, suggesting that the loading tasks did not effectively interfere with the left key pressing task. As can be expected then, no ITC-SOPI differences exist between conditions, forcing the conclusion that neither attention nor working memory load tasks reduce presence below the baseline condition. However, the analysis modeling left key presses from presence shows an effect – namely, spatial and naturalness factor scores predict task performance on left key presses (regardless of task condition). The direction of correlations is however somewhat counterintuitive. The spatial factor

has a positive correlation with task performance as can be expected: subjects who experienced the video as a space would have been able to predict the doorway events better, and would therefore have been able to press the left key even if their attention was diverted to the secondary task. This explanation is in line with other explanations of the relationship between presence and task performance, particularly those that invoke situational awareness (Bystrom *et al.*, 1999a; Laptaned, 2006; Youngblut & Perrin, 2002), and replicates the task-presence relationship found by Slater *et al.* (1996).

The negative relationship between naturalness and task performance is puzzling. It implies that finding the experience realistic and having it match expectations about the environment impedes task performance. It is very difficult to explain this result, as ITC-SOPI scores are almost always positively correlated to each other (Lessiter *et al.*, 2001), and the CLCC model was built with this assumption in mind. This break from the usual pattern of correlations might indicate why studies that use single-factor presence measures find the task performance-presence relationship so difficult to replicate (Laptaned, 2006; for instance Sas & O'Hare, 2001; Sas *et al.*, 2004). If a single measure is conflating spatial presence and naturalness (which in this study had similar strength but inverted effects), then it is likely that those effects would cancel out and show no difference in the measure.

Although somewhat circuitous, the correlation between ITC-SOPI factors (spatial and naturalness) and task performance in this study, does have important meaning in terms of this CLCC prediction. As discussed in 7.8.4 above, performance on tasks which can theoretically show interference effects can be used as an indirect measure of processing capacity use; and the relationship between some aspects of presence and performance on this task may indicate that presence also makes use of these resources. Further work, particularly with tasks which show an independent difference across interference conditions, will be needed to clarify this relationship.

2. *A working memory loading task will reduce presence more than an attention loading task* – The analysis of right key shows fewer right key presses in the

working memory condition than in the attention condition, suggesting that subjects in the working memory condition were more likely to miss an event – this in turn shows that the task manipulation was successful. It should be noted that there was no significant difference between these two conditions on left key presses. This indicates a potential weakness in the design –the two right key press tasks (press on walking through a door way as compared to press on seeing the blue block flash on the video) may have been of differing difficulty. Given that the subjects did not know that the left video was the VE of interest (or that the study was about presence or environments), it is a difficult result to explain. One possibility is that the effect size of the left key press task is smaller than that of the right key press task, and so a sample size of 15 per cell is gives enough power to find the right key press effect, but not the left key press effect. Be that as it may, more important is the lack of ITC-SOPI differences across conditions. Although the task manipulation was partly successful, the conditions produced no significant differences in any of the four ITC-SOPI factors, suggesting that the scores produced by the working memory load and attention conditions are the same. This finding echoes the pattern of findings from Studies 1 and 2, which similarly showed that working memory manipulations led to no interferences on presence scores.

7.10 Conclusion

This study generally corroborates the findings of Studies 1 and 2, and shows that working memory and attention manipulations do not generally interfere with presence, although there were some small effects (presence was able to predict one aspect of task performance, indicating some link between resource allocation and presence). However, the study was able to find no difference between the effects of attention load and working memory load on presence. This contradicts the predictions of presence models which propose an important role for attention in presence, including the CLCC model, the environment selection model (Slater, 2002), the three-pole model (Biocca, 2003), the FLS model (Waterworth & Waterworth, 2001) and the MEC model (Wirth *et al.*, 2007). This study again suggests (as did Studies 1 and 2) that the role of processing capacity or attention in presence is either very small, or only discernible across large manipulations (such as when the secondary interference task is extremely difficult). Although it is worth following up these leads with further research, the pattern of findings form the all three of these studies (together with the findings of S. Lee *et al.*, 2004a which also similarly failed to find an attention effect on presence) suggest that if the study is to have any chance of finding an effect, a large scale manipulation and large samples should be used, as the effect is likely to be very small.

Chapter 8: Study 4 – An exploration of experience effects and cognitive processing in presence

The previous three studies examined one of the two major innovations of the CLCC model - the role of working memory in presence. The next three studies examine the second innovation of the model; namely, the role of declarative memory and semantic bias in presence. Studying the role of working memory is the easier of the two problems, because manipulating working memory load experimentally is simple. However, semantic effects are harder to pin down, because they are bound to particular content areas. One cannot simply measure how much a subject knows; one must declare a content area and measure how much they know about *that area* (Nichols & Stich, 2000). Similarly, one cannot simply examine semantic bias without considering the actual content of the bias, because different processing strategies may be used for different types of data; for instance, a subject may be an expert in one narrow domain (Ericsson & Lehmann, 1996), or they may be processing data for which specialized neural circuits exist (Plotkin, 1998; Sperber & Hirschfeld, 1999).

8.1 Studying content effects on presence

To limit the range of content domains and processing strategies, the current study makes use of a specific population - computer gamers. Gamers are an interesting population for presence research, for a number of reasons. First, they are probably the only large population which regularly experiences presence in interactive systems. Second, many gamers have been playing for years, which allows the examination of possible learning and experience effects. Third, although there are a wide range of games available, only a small number of genres exist (first-person shooters, real-time strategy, etc), which limits the range of cognitive strategies and knowledge associated with the activity. Fourth, unlike the often tedious tasks used in research VEs, computer games are designed to capture and hold user's attention and interest, which means that gamers often have experienced intense presence. Finally, being a computer literate group, they are an easily accessible population. There are some problems associated with using this group also - particularly the significant gender inbalance (see 8.3 below) and the lack of control over the nature and extent of previous experiences. Using gamers can thus provide a high degree of ecological validity, while trading off some internal validity.

8.1.1 Time, experience and self-rated importance of presence

Experimental studies have found relationships between presence and various time and experience related factors such as age (e.g. Youngblut & Perrin, 2002), game playing experience (e.g. Usoh *et al.*, 1999) and previous exposure to VEs (e.g. Meehan *et al.*, 2002). While such experiments provide valuable insights, their ecological validity is limited by the types of tasks, environments and often contrived experimental conditions they use (this limit in ecological validity is of course the case for experiments in general - Rosenthal & Rosnow, 1991). The current study will examine the "average" degree of presence during game playing in a general sense in an effort to increase the generalizability of the findings. Rather than examine one particular experience, this study uses the somewhat unusual strategy of asking subjects to rate *how important they consider presence to be to their gaming* in general (this is termed *self-rated presence importance*). The reasoning is that if subjects have enjoyable or compelling presence experiences while gaming, their ratings of presence importance will be higher. This measure also gets around problems associated with measuring presence experiences across many different software platforms, which would not be directly comparable, as they would all produce differing levels of presence by immersion effects alone. This variable is measured by means of self-report items such as “A game should make me feel as if I am transported to inside the game world.” and “I prefer games which create a sense of being in a place” (see 8.5 below). Such items are straightforward to report on, and can be accurately responded to. Due to this choice of variable, this study does not make use of standard presence questionnaires (these are designed to measure presence at one particular experience). It should be noted that *self-rated presence importance* is a construct distinct from presence itself, although, as argued above, it is highly likely to be a predictor of presence. While this choice of variable reduces the reliability of the measure (as the inter-item correlations in the measure are unknown - Anastasi & Urbina, 1996), the potential increase in ecological validity balances this risk sufficiently to justify its use in an exploratory study.

8.1.2 Presence maximization strategies

When examining learning effects, one must consider the possibility that gamers may have learnt or evolved strategies to maximize their presence experiences (this is most likely for gamers who rate presence as being an important part of the gaming experience). If such strategies exist, then their use is likely to vary with time-related factors (length of time playing, age, etc), because more experienced subjects are more likely to have discovered them, and also, if they are effective, more experienced players will have had their use more strongly reinforced. The literature reveals two possible presence maximization strategies which the average gamer could implement: minimizing distracters to maintain attention focused on the display (Slater, 2002; J. A. Waterworth & E. L. Waterworth, 2003b; Wirth *et al.*, 2007), and improving display fidelity by maintaining up-to-date computer equipment (IJsselsteijn *et al.*, 2000; Slater, 2003a; Witmer *et al.*, 2005).

8.1.3 Cognitive processing

While presence management includes factors associated with immersion and bottom-up factors, top-down processes are represented in this study via two cognitive processes which are of primary importance in the CLCC model: thematic inertia and the integration of information based on semantic meaning:

Thematic inertia - this is a measure of the extent of semantic networks in declarative memory, and how long it takes for activation to decay from them. A subject who is experienced in a particular content area will have wider networks, and therefore more thematic inertia (see 4.3.10 in chapter 4 for a discussion). According to the CLCC model, thematic inertia is a complex predictor of presence. A wider semantic network will lead to more specific expectations about the VE, and if these expectations are matched by the display, a more compelling presence experience will result; however, if these experiences are not matched by the display, the presence experience can be compromised. In this study, thematic inertia is operationalized as the tendency for a subject to engage in thematically related activities (e.g. reading about aviation, as well as playing aviation related games). Specifically, it includes situations where non-simulation activities (reading a book, taking a real flight) led to either a desire to play a flight simulator, or the actual playing of a flight simulator.

Integration of information by semantic meaning – Nunez & Blake (2003a) showed that providing subjects with semantically related information before VE exposure could lead to an increase in presence under particular display conditions. In the current study, this effect will the explored by examining the *capacity to integrate non-diegetic information*. In film studies, the term 'non-diegetic' refers to information which does not emanate from the story world, but from some point external to the film world (for instance, background music or narration - Bordwell, 1989). Non-diegetic information is thus coherent with the environment in meaning, but not in terms of fidelity. The CLCC model predicts that presence is associated with the extent of semantic meaning which the subject attaches to the environment; therefore, any information which provides more detail about the environment (even if non-diegetic) should contribute to the presence experience, even if it reduces the fidelity of the display.

8.2 CLCC model predictions

As this is an exploratory study, it is not possible to generate hard predictions. However, one can state general expectations based on the CLCC model:

Presence will be positively correlated with game-playing experience – this expectation comes from two sources in the model. First, given that most computer games involve a complex user-interface, it is highly likely that there exists a learned media decoder to process the gaming display. This means that one can expect experienced gamers to be more efficient at decoding the display, and therefore have more compelling presence experiences (although the evidence from Study 2 in chapter 6 suggests that this should only be a minor factor). Second, experienced gamers will have learned the genre related conventions and iconography (stored in declarative memory) which helps them associate their existing semantic knowledge to the elements of the display. This means that they will more easily chunk content in working memory, and their interactions with the game will be more meaningful, and thus lead to more presence.

More extensive knowledge networks (as indicated by higher thematic inertia) will lead to a lower degree of presence – gamers who are expert in the game's content area will have more extensive semantic knowledge networks; therefore, they will therefore more thematic inertia, and more specific expectations about game content. If the game is not able to match these expectations, the presence experience will suffer. Given that all games have a limited capacity to match highly specific expectations, gamers who are content experts will usually have unmatched expectations, and therefore reduced presence. Note that this prediction does not contradict prediction 1 above. Prediction 1 refers to generic game playing experience, and this prediction refers to specific content-area related knowledge. For instance, if a World War II based game is given to a group of gamers, then one will expect those who have been gaming longer to have more significant presence experiences, due to their mastery of the interface, and their ability to understand the display; however, those gamers who know more about World War II as a semantic knowledge domain will have reduced presence experiences due to unmatched expectations. Note also that gaming experience and content domain knowledge are not necessarily correlated, as one can be a content expert without having ever played a game, as well as being an experienced gamer, without knowing very much about one particular content area.

Presence is associated with the integration of content related semantic information – recall that the temporary structures in working memory act as connections between sensory input and semantic concepts. These structures associate information only based on its decoded semantic meaning, and not the source from which it arises. Therefore, one can expect that any information which is semantically related to the temporary structure will be incorporated into it, regardless of its source. As more extensive temporary structures are associated with higher levels of presence, any semantically related information (even if non-diegetic) which is incorporated into the temporary structure should increase the level of presence.

8.3 Sample

101 responses were collected over a one-week period (see 8.4 below for a description of the procedure). The mean age was 22.13 years (S=3.23), with a minimum of 17 and a maximum of 34. Only 3 respondents were women (2.97%). This seems a gross overrepresentation of men, if one compares to widely quoted North American data collected in 2000 by the Entertainment Software Association, which found 43% of their gamers sample to be women (Entertainment.Software.Association, 2000). However, the proportion of women players seems to vary according to many variables, such as game type. Avsim.com, a major web-site associated with flight simulation gaming, in their 2003 survey of 14,247 flight-simulation gamers, found that only 2.6% of the sample were women (avsim.com, 2003). It is likely that the proportion of women gamers also varies with cultural variables. In South Africa, where this study was conducted, a study of young women found as few as 13% who played computer games (Sander & Galpin, 1994). Therefore, while the current sample clearly under-represents women, it is not clear by how much.

8.4 Procedure

The study used a survey, which was preferred over an experiment for three reasons. First, to create an experiment to study long-term experience effects would require a lengthy longitudinal study. A survey allows one to examine data from subjects who have, in most cases, several years of experience, which would be practically impossible to recreate under controlled conditions. Second, an experiment must be limited in terms of the virtual environments or games used during manipulation, while a survey allows for subjects to have gained their experience with a wide range of environments and games; and third, surveys provide a great degree of ecological validity provided they are given a large and diverse enough sample (Rosenthal & Rosnow, 1991). Given that this study is exploratory, it was felt that the loss of experimental control was balanced by the gain in ecological validity. The loss of internal validity can also be overcome by verifying interesting findings in a follow-up study (see studies 5 and 6 in chapters 9 and 10 for these follow-ups).

The study was advertised as a 'computer gaming habits survey' at a South African university, and the survey was posted on a web-site. When subjects arrived at the web-site, they were given a brief description of the study (its length and form), and given a link which started the study (clicking on the link indicated consent to participate). All 40 items in the study were presented on one page, with a single item per line. The survey page was structured such that the subject could not submit their data without first having completed all of the items on the page. Once the subjects submitted their data, they were thanked for their participation. The study was left open for a one-week period in the middle of the teaching term.

8.4.1 Categorization of game types

For this study, computer games were broadly divided computer games into two categories: those which aim to produce presence (‘presence games’) and those which do not (‘non-presence games’). Presence games include among others simulators, role-playing games and first-person shooters, while non-presence games include real-time strategy, abstract puzzles and fighting games. This division was made so as to allow the separation of effects associated with gaming and interactive entertainment in general from effects which are unique to presence.

8.5 Measures

A 40-item instrument measuring 10 factors was created for this study. 6 of these factors were time and experience related factors (Table 8.1 below), and the other 4 were cognitive factors (Table 8.2 below). For all items except those measuring time (e.g. length of time playing presence games or age), a seven point Likert-type response format was used. The full questionnaire can be found in appendix C.

The factors used were:

Length of time playing presence games –the length of time (in years) which the subject has been playing presence games (see 8.4.1 above). This factor allows measurement of slow forming strategies and effects.

Frequency of playing presence games –how often in a typical week the subject plays presence games. This factor allows the estimation of the short-term contribution of presence experiences to the evolution of presence maximization strategies. It also allows an estimation of the degree to which thematic inertia might aid in the development of such strategies, as a frequent gamer experiences higher levels of VE relevant activation more often, which may lead to an increase in thematic inertia.

Frequency of playing non-presence games – this is a control for effects which may be related to using interactive entertainment as opposed to presence itself. If a particular effect is related to gaming in general (as opposed to specifically presence gaming), then this factor will become a significant predictor. It also acts as a control for general computer experience effects, as it effectively measures the media consumption of the subject.

Knowledge of computers – this is included as a control following the advice of Lessiter *et al* (2001), who argue that experience with and knowledge of the medium being used can alter the presence experience for the subject.

Knowledge of games – this also follows Lessiter *et al*’s (2001) arguments, but is more specific to gaming as opposed to computers in general.

Age – this is a general control for maturation or generational effects, following the correlation between age and presence reported by Barfield & Weghorst (1993).

Factor	Number of items	Example Item
Length of time playing presence games	3	How long have you been playing first person shooters?
Frequency of playing presence games	3	How often do you play simulators?
Frequency of playing non-presence games	3	How often do you play fighting games?
Knowledge of computers	1	How much knowledge do you have about how computers work?
Knowledge of games	1	How much knowledge do you have about how computer games works?
Age	1	Your age:

Table 8.1: Time-related and experience factors

Factor	Number of items	Example Item
Integration of non-diegetic information	5	Inappropriate music in a game can ruin the game experience for me.
Self-rated importance of presence	6	A game should make me feel as if I am transported to inside the game world.
Thematic inertia	6	After watching a TV program or film, I often feel like playing a game that is similar to the film or program.
Presence maximization	6	When I play, I turn off the lights and try to keep the room dark.

Table 8.2: Cognitive factors

Integration of non-diegetic information – this is an estimate of the degree to which the subject is integrates thematically related information, regardless of source or modality. It reflects the action of the CLCC model's knowledge clusters, which associate information based only on semantic content, and not origin. This factor focuses specifically on non-diegetic information for two reasons: first, the effects of multimodality (i.e. the case where all information is diegetic) is already well studied in the field (see Bystrom & Barfield, 1999; Hoffman *et al.*, 1999; Sallnäs, 1999 for examples; and section 3.3.1.3 in chapter 3 for a discussion of these findings). Second, non-diegetic information clusters with diegetic information purely by its semantic content, making it a top-down effect, which is a relatively unexplored topic in presence.

Self-rated importance of presence – this measures how important the subject finds the presence experience to their gaming. It is hypothesized that subjects who have well-developed capacities to experience presence will score highly on this factor in the context of presence games, while the context of non-presence games will not lead to an effect. Subjects who do not have the capacity to experience presence should score low on this factor regardless of game type, as their lack of presence experiences would not have lead to them experiencing any benefits during gaming.

Thematic inertia – in the CLCC model, thematic inertia represents the extent to which activation spreads among active knowledge clusters in declarative memory (see 4.3.10 in chapter 4). Active knowledge clusters remain active after stimulation has ceased, and this results in a desire to continue with related activities (until the cluster becomes inactive). In the survey, this is a operationalized as the degree to which engaging in some activity leads to a desire to engage in another activity which is related by theme or semantic content; for example, reading a science-fiction novel leading to a desire to play a science-fiction game.

Presence maximization – this estimates the extent to which subjects control their environment to maximize the probability of experiencing presence. This factor captures this in terms of two strategies which the literature argues correlate with presence: reducing possible distractions during the experience (Slater & Steed, 2000; J. A. Waterworth & E. L. Waterworth, 2003b; Wirth *et al.*, 2007), and maintaining up-to-date computer hardware and software to maximize fidelity and immersion (Barfield *et al.*, 1998; IJsselsteijn *et al.*, 2000; Witmer *et al.*, 2005). Subjects who score high on this factor are therefore expected to be more likely to experience presence when playing presence games.

8.6 Analysis & results

As all the variables used in the study were continuous, the data were analyzed using a series of linear regression models (Neter *et al.*, 1988). In keeping with the evolutionary approach of this study, the variable modeled was either the *presence maximization* factor or the *self-rated importance of presence* factor, which would indicate to what degree subjects had developed the ability to structure their experiences for presence, and the degree to which those efforts were successful.

8.6.1 Cognitive factors and time/experience effects

As the rate of schemata activation and decay is probably set at an early age and changes little over time (Rumelhart & Ortony, 1977) it was expected that thematic inertia would not be predicted by the time and experience factors. A multiple regression analysis with the six time factors as predictors and thematic inertia as the criterion confirms this prediction: $F = 0.89$; df= (6, 66); $p < 0.505$. The lack of effect on this regression demonstrates the independence thematic inertia from learning factors.

In the CLCC model, the integration of non-diegetic information is a function of the media decoder, which is a learned structure (see 4.3.5 in chapter 4). If this is true, then one can expect experience effects. To investigate this possibility, a multiple regression to predict *capacity to integrate non-diegetic information* from the time factors was computed, and was indeed significant: $F = 2.42$; df = (6, 66); $p < 0.036$; $R^2 = 0.18$. Of the six time factors, only *length of time playing presence games* is significant (see

Table 8.3). Although this result can be interpreted as supporting a media-decoder based 'learning to decode' hypothesis, it is also possible that those subjects who are better able to integrate non-diegetic information tend to have a more enjoyable presence experience during gaming and thus keep playing this type of game for longer periods; this second interpretation might seem to be more in keeping with the lack of evidence for separate media decoders found in Study 2 (see chapter 6).

8.6.2 Cognitive factors as predictors of self-rated presence importance

The two cognitive factors (thematic inertia and the capacity to integrate non-diegetic information) show a significant correlation with each other ($r = 0.36$; $n = 101$; $p < 0.01$). This supports the notion that they share a common cognitive basis, as predicted by the CLCC model. To determine if these cognitive factors are related to presence, they were used as predictors of *self-rated presence importance* in a multiple regression analysis. This gives a significant model: $F = 12.49$; $df = (2, 98)$; $p < 0.0001$; $R^2 = .202$. In this regression, only integration of non-diegetic information is a significant predictor, although thematic inertia shows a large effect, and could be considered as significant (see Table 8.4 below). When the effect of thematic inertia on self-rated presence importance is examined on an item-by-item basis (controlling for the integration of non-diegetic information), thematic inertia was found to be a significant predictor of two items: "I prefer games which create a sense of being in a place." (partial $r = 0.29$; $p < 0.016$; $R^2 = 0.16$) and "For me, the most important aspect of game playing is the ability to explore other worlds." (partial $r = 0.25$; $p < 0.022$; $R^2 = 0.16$).

Predictor	**Partial r**	***p***
Length of time playing presence games	0.365	0.002
Frequency of playing presence games	0.108	0.379
Frequency of playing non-presence games	-0.088	0.475
Knowledge of computers	-0.030	0.805
Knowledge of games	-0.028	0.818
Age	-0.066	0.588

Table 8.3: Partial correlations for the model predicting *integration of non-diegetic information* from the six time and experience related factors. Significant predictors ($p < 0.05$) are highlighted.

Predictor	Partial r	*p*
Integration of non-diegetic information	0.339	0.0005
Thematic inertia	0.195	0.051

Table 8.4: Partial correlations for the model predicting *self-rated presence importance* from the cognitive factors. Significant predictors (p < 0.05) are highlighted.

8.6.3 Experience effects on self-rated presence importance

If presence is predicted by particular learned cognitive processes as suggested by the CLCC model, then more experienced players should experience more presence. A regression model to predict the self-rated importance of presence in games using all six time-related factors as predictors is indeed significant: F = 2.78; df = (6, 66); $p < 0.017$; $R^2 = 0.202$. An examination of the partial correlations to control for inter-variable dependencies shows the only significant predictor to be *frequency of presence game playing* (see Table 8.5).

Predictor	M	S.D	Partial r	*p*
Length of time playing presence games (years)	3.05	0.544	0.064	0.603
Frequency of playing presence games (times per week)	1.05	0.53	0.351	0.003
Frequency of playing non-presence games (times per week)	0.98	0.51	-0.220	0.073
Knowledge of computers (1 = novice, 7 = expert)	4.5	0.71	-0.155	0.204
Knowledge of games (1 = novice, 7 = expert)	3.13	0.67	-0.135	0.269
Age (years)	22.13	3.2	-0.114	0.354

Table 8.5: Partial correlations for the model predicting *self-rated presence importance* from the six time and experience related factors. Significant predictors (p < 0.05) are highlighted.

An examination of each of the six items composing the self-rated importance of presence factor shows that one item ("The quality of a game's sounds are very important for my game experience.") was also inversely predicted by *frequency of non-presence game playing* (partial r = -0.25; t(66) = -2.11; $p < 0.037$). Only one item ("For me, the most important aspect of game playing is the ability to explore other worlds."), was not predicted by time-related factors at all. The lack of effect on this item is probably attributable to the wording of the item. Although some players may enjoy exploring game worlds (a high-presence activity), most games make exploration a secondary activity – the player's primary, and therefore most important, goals (winning a fight, solving a puzzle, etc.) are often non-presence activities.

8.6.4 Effectiveness of presence maximization strategies

If the basic evolutionary assumption in this study is correct, then the degree to which subjects engaged in presence maximization strategies should predict their self-rated importance of presence. This was tested using a simple linear regression model using *presence maximization* as the predictor for *self-rated presence importance*. The subsequent model was, as expected, significant, although it explained only a small amount of variance (F=18.87; df=1, 99; p<0.0005; $R^2 = 0.15$).

An item-by-item investigation of the *self-rated presence importance* factor showed that only two of the six items in the factor failed to show this relationship. The items "The quality of a game's sounds are very important for my game experience." and "I prefer games which create a sense of being in a place" were not predicted by presence maximization strategies. This may indicate that these items tap into inherent properties of the subjects, and are therefore not subject to experience effects.

8.6.5 Learning to maximize presence

If presence maximization strategies are indeed evolved using presence as a feedback mechanism, then one would expect the most experienced players to make the most use of these strategies. Again, a multiple regression was computed with all six time related factors as predictors for presence maximization strategies: F = 2.83; df = (6,66); $p < 0.016$; $R^2 = 0.204$. Only *knowledge of computer games* was a significant predictor in this model (see Table 8.6 below). Interestingly, the partial correlation shows that higher knowledge of game workings is associated with *reduced* efforts to manage presence.

This finding suggests that gamers who understand games more (and presumably the reliance of modern games on specialized computer hardware) would at least make an effort to keep their equipment up to date. A possible confound in this study may have been that maintaining updated computer equipment was beyond the economic reach of our sample of university students. Evidence for this conjecture was found when comparing the two items "As far as I can afford it, I make sure my computer has the best hardware for playing games." and "I will consider upgrading my computer to play a particular game." How long players had been playing presence games was indeed a significant predictor for the second item, which is hypothetical and thus not bounded by practicalities like the subject's bank balance (partial r = 0.28; t(66) = 2.39; $p < 0.019$), but not for the first, which represents what the subject actually can do.

Predictor	**Partial r**	***p***
Length of time playing presence games	0.196	0.108
Frequency of playing presence games	0.150	0.221
Frequency of playing non-presence games	-0.093	0.448
Knowledge of computers	-0.015	0.902
Knowledge of games	-0.311	0.009
Age	-0.163	0.181

Table 8.6: Partial correlations for the model predicting *presence maximization strategies* from the six time and experience related factors. Significant predictors ($p < 0.05$) are highlighted.

For the distraction related items, there were indications that time-related factors play a role. For the item "If I am disturbed while I am playing, it ruins the experience for me.", both age (partial r = 0.25; $p < 0.043$) and how long the player had been playing presence games (partial r = 0.29; $p < 0.014$) were significant predictors (note that although both these predictors are likely to co-vary, the partial correlation has removed the shared variance between them).

8.7 Discussion

8.7.1 Support for CLCC predictions

Although this study could only generate general expectations rather than definite predictions, it is worth considering the evidence as a beginning point for discussion:

Presence will be positively correlated with game-playing experience – this was generally supported. When modeling self-rated presence importance, the frequency of presence-game playing was a significant predictor. However, the length of time playing presence games (in years) was not a predictor. This may indicate that the presence learning effect is a medium-term effect, and not a long-term effect (that is, playing games often keeps the effect active, but it does not accumulate over years of experience – see 8.7.2 below). Note that this is not simply a general game-playing effect; frequency of playing non-presence games was not a predictor, indicating that it is the frequency of presence experiences which produces the effect.

More extensive knowledge networks (as indicated by higher thematic inertia) will lead to a lower degree of presence – in the model of self-rated presence importance predicted by cognitive factors, thematic inertia was a predictor of borderline significance ($p < 0.051$). This suggests that thematic inertia is probably related to self-rated presence importance, as suggested by the CLCC model, although it is likely a weak effect. The item-by-item examination of self-rated importance of presence indicates that thematic inertia is a likely predictor for spatial presence rather than the cognitively higher forms of presence (such as engagement and naturalness).

Presence is associated with the integration of semantic information – this expectation was clearly supported by the model of self-rated presence importance predicted form cognitive factors – the integration of non-diegetic information is a significant predictor.

8.7.2 Learning and experience effects

This is only a relational study and cannot show causation; nonetheless, the data show interesting trends with regard to experience effects and cognition in presence. First, it seems that the most reliable experience related predictor of self rated presence importance is the proportion of gaming time spent playing presence games. This suggests that presence displays a mid-term, slow-decay effect: it begins by one presence experience leading to a state which provides some thematic inertia, and a latent positive benefit for the next presence experience. If no presence experience occurs for a while, the latent benefit decays and disappears (this is indicated by the fact that while *frequency* of presence game playing is positively associated with presence, *length of time* having played presence games does not). A competing explanation is that for some individuals, the presence experience becomes highly desirable, and so they seek it often, leading to higher gaming frequency. However, as the CLCC model provides no explanation for these individual differences in the desirability of presence experiences, the former explanation is preferred.

The data do not seem to indicate that users become desensitized to presence. This is inferred from the general lack of effect of the length of time playing presence games. Indeed, the opposite may be true, as age has a weak positive effect on self-rated presence importance (a finding which is supported by the positive correlation found by, among others, Youngblut & Perrin, 2002).

8.7.3 Presence maximization strategies

Caution should be taken in the interpretation of the effect of these presence maximization strategies. As the design used in this study is a relational one, it is also possible that the importance give by gamers to presence could lead them to engage in presence maximization strategies (rather than the reversed interpretation made above). Nonetheless, the data suggest that gamers do successfully engage in strategies to maximize their presence. Interestingly, these efforts generally vary (inversely) with knowledge of how games work. There are two possible explanations for this phenomenon: one is that gamers' knowledge about the technical aspects of the game interferes with their ability to suspend their disbelief during play; either by occupying working memory, or by focusing their attention on the display rather than the virtual environment. The second is that gamers have naïve theories of how presence 'works', but more experienced gamers (who probably obtain most of their knowledge from gaming websites and gaming magazines) believe the common game marketing line that the software is largely responsible for presence, and thus make no effort to control their own environment during play. It would be necessary to explicitly tap into these naïve theories to validate this hypothesis. Regardless of what gamers believe about their presence experiences, it seems that the presence maximization techniques do have some effect, although with little consistency.

These findings may be partly obscured by economic factors. One of the two presence maximization strategies measured was the maintaining computer hardware that was up-to-date. It is likely that the gamers in our sample want buy the newest hardware, but as almost all were university students, their economic realities would interfere with this goal. Evidence for this comes from the comparison of two items in the factor: the first measures real money expenditure (in which no time effect was found, contrary to expectation), and the second measures hypothetical expenditure (for which length of time playing presence games was a predictor, as expected). This implies that long-time players of presence games recognize the importance of maintaining updated hardware, but may not always be capable of doing so in practice.

8.7.4 The role of sound

An interesting finding about the importance of sound is worth mentioning. Gamers' ratings of the importance of sound to the game experience were not associated with presence management strategies, but their ratings of the importance of graphics were. This finding is interesting because the literature does not make a distinction between modalities in terms of their contribution to presence. This could, of course, simply be a deficit in the literature: our definition of presence maximization strategies is based on reported factors which have a high degree of empirical validation; and it is currently the case that the visual modality has received far more research attention than other modalities. Therefore, it may be the case that presence maximization strategies related to sound do exist, but the literature has not explored what they are.

The importance of sound was however strongly linked to frequency of presence game playing. This may imply that the integration of sound into the presence experience is not affected by a player's efforts, but does improve with repeated exposure. This may suggest that the contribution of sound to presence is processed separately from other modalities, contrary to the CLCC expectation that all sound is decoded and stripped of source before further processing. A more likely explanation for this finding is that it arises from an artifact of the study's method, already discussed above. As there is very little research on specific contributions of sound to presence (as compared to similar research on the contribution of graphics, for example), very few sound relevant presence management strategies were identified from the literature and included in the survey. Therefore, it might be the case that gamers do indeed engage in some strategies, which were unidentified by this study.

8.8 Conclusion

The findings in this study are useful in providing some validation for the top-down aspects of the CLCC model, but there are some weaknesses that must be addressed in a follow-up (many of these are addressed in study 5, see chapter 9):
The study consciously avoided any one particular content area to maintain generalizability. This strategy however increases the amount of error variance in the study, by lacking a measure of the actual degree of subjects' content knowledge.

The study supports the interesting notion that general and specific expectations about the VE may have different effects on presence. The measure used, however, made no strong distinction between general and specific expectations. Part of this problem arises because it is not possible to measure the degree of expectation without focusing on one particular content area (see 1 above).

The measure of presence used was too indirect to provide conclusive support for the CLCC model. Also, as a standard measure and concept of presence was not used, it is difficult to compare these findings with published studies in the literature.
These weaknesses aside, the study did confirm two important aspects of the CLCC model. First, it confirmed the importance of declarative memory in presence by showing learning and experience effects on self-rated importance of presence. Second, it showed the importance of complex, semantic information processing by showing that the integration of non-diegetic information also has a potential effect on self-rated importance of presence.

Chapter 9: Study 5 – The role of content knowledge in presence

Study 4 was an exploration of various content and experience effects on presence. Two of the results of that study are notable: that content knowledge (indicated by learning and thematic inertia) is associated with presence, and that cognitive processing (indicated by the integration of non-diegetic information) affects presence. The exploratory nature of that study prevented specific predictions (see section 8.8 in chapter 8). The present study aims to verify these findings, and overcome some of the weaknesses of Study 4.

9.1.1 Theories of content effects on presence

There is controversy surrounding the role of content in presence. Slater (2003a) explicitly separates spatial presence (when a virtual place is experienced as if it were not mediated), from other concepts such as engagement and involvement, which are regarded by others (for instance, Lessiter *et al.*, 2001) as being a part of presence. He argues by analogy that presence occurs by the *form* in which information is presented to the subject. Interest and involvement are brought about by the subject's relationship to the content. He thus concludes that content related factors do not determine presence. In general the literature does not agree with this distinction between spatial presence and other concepts such as engagement – see section 2.4.1 in chapter 2 and section 4.6 in chapter 4 for more detailed discussions of the relationship between these concepts.

In media theory, form and content are generally considered separate concepts (Calvert, 2001). Logically, they are neither equivalent nor necessarily related. However, this belief exists without much empirical evidence. In the presence literature, the role of content has been mostly discussed theoretically only (for instance, Fencott, 1999; Nunez & Blake, 2001; Wirth *et al.*, 2007). The consensus seems to be that presence requires the environment must make sense or carry meaning for the subject; and it is the content that provides meaning. Some support exists for this idea. A factor analysis of eight presence measures, (Schubert *et al.*, 2001) found that *drama* (the degree to which the VE presents a story in which events unfold in a meaningful, predictable way) ranked 4th out of 8 extracted factors, with an eigenvalue greater than 3.

9.1.2 A theory of content in presence

One reason why content may not have attracted much research attention is the difficulty of operationalizing it as a variable. Are there dimensions along which content can be measured? Can the impact of a particular content area be measured? It is unlikely that particular content areas will have presence effects. More likely, content effects arise from how the data in the VE interacts with the knowledge held by the subject. This interaction can be modeled in terms of user expectations. As the user begins the experience, the VE content slightly activates some knowledge clusters in declarative memory. That activation then spreads though the knowledge cluster, leading to expectations of subsequent experiences in the VE (see sections 4.3.6. and 4.3.10 in chapter 4). If the VE matches these expectations, one can expect a coherent construction of the VE and an enhanced presence experience (particularly for cognitively higher forms of presence such as engagement and naturalness).

If this is the case, then one can categorize users' knowledge in terms of the types of expectations they will produce. Detailed knowledge will produce highly specific expectations, which will be hard for a VE to match; general knowledge will lead diffuse expectations which are easy for a VE system to match (see prediction 2 in section 8.2 in chapter 8). Therefore, experts should find VEs of their content areas to be largely unsatisfactory (unless the simulation has been designed to an extremely high degree of content fidelity), as they would notice errors and inaccuracies. On the other hand, a novice would find the same simulation satisfactory due to having only very general expectations for that content, and might therefore experience more presence due to fewer processing interruptions. This is analogous to the well-known "uncanny valley" found in simulations of humans (Mori, 1970). Almost all subjects are experts in processing the human form human form (albeit largely implicitly), which produces very specific expectations. A simulation of a human must be of an extremely high degree of fidelity to match such an expectation; indeed, most contemporary systems fail at this task, leaving users largely unsatisfied by the simulation. The CLCC model's explanation of the interaction between knowledge, expectation, and the matching of these expectations can essentially be understood as a general explanation of the uncanny valley phenomenon, from a cognitive perspective.

9.2 Predictions of content effects made by the CLCC model

The CLCC model makes the following predictions about knowledge and expectation of VE content:

1. *Generalized knowledge of content will increase the presence experience, while specific knowledge will reduce the presence experience* – This arises from Study 4 (see section 8.8 in chapter 8), which hypothesized that specific and general knowledge of content may have different effects on presence. According to the CLCC model, content knowledge is represented as networks of knowledge clusters in declarative memory (see section 4.3.6 in chapter 4). When one of these knowledge clusters becomes active, it acts as a context or expectation, which the data arriving bottom-up from the stimulus attenuator content can match or fail to match. A match leads to an increase in the semantic bias, a more coherent construction of the VE, and a high degree of presence (particularly for cognitively higher forms of presence such as engagement and naturalness - see section 4.6 in chapter 4). Specific knowledge implies a wide, well-connected network of knowledge clusters, which include a large number of details connected to each concept (hence allowing encoding highly specific knowledge), which would lead to extensive expectations, which bottom-up data would be less likely to match, and thus one predicts a reduced presence experience. On the other hand, generalized knowledge is associated with a small network of knowledge clusters, with few details associated with each concept, which is more likely to be matched by bottom-up data, and therefore an increased presence experience.

2. *High level semantic processing contributes to presence at least as much as perceptual processing* – The CLCC model proposes that presence occurs due to the formation of temporary structures in working memory which have a bias towards processing the virtual environment (see section 4.3.3 in chapter 4). These temporary structures arise due to the interaction of perceptual data

arriving from the stimulus attenuator, and semantic information arising from active knowledge clusters. The model gives no preference to either of these sources in this process, although it does predict that semantic processing will have more impact on the cognitively more complex forms of presence such as engagement and naturalness. One can therefore predict that semantic processing should play an equal role to perceptual processing in the presence experience, particularly for the cognitively higher forms of presence.

9.3 Procedure

As in Study 4 (see chapter 8), a relational design was preferred in order to collect a large sample of habitual VE users. The study took the form of an online survey of flight simulation game players (advertised as a 'flight simulator gaming habits' study). Seven web-sites were selected to advertise the study: these were either web-portals to the flight simulation community, or web forum sites which primarily discuss flight simulation. An incentive for participation was offered: Flight1.com, an on-line retailer of flight simulation products, was recruited as a sponsor of the study. They provided three popular flight simulator products as prizes for a random draw of subjects.

Subjects were provided with information about the study (enough to meet the informed consent ethical requirement while minimizing expectancy effects), and if they agreed to participate, were presented with the questionnaires. The entire survey was divided into five sections to improve page download speed. Subjects could only continue to the next section of the study once they had completed all the items on the current section. To ensure that subjects did not feel coerced into participating by this mechanism, the header of each page held a reminder that they could exit the study at any time by simply closing the browser window. Subjects were first presented with the ten content, cognitive and control factors (see Table 9.1 below), and then were asked to report the title of the last flight simulator they played. They were then asked to complete the ITC-SOPI with regard to that last session. With the exception of not administering the ITC-SOPI immediately after the VE experience, all the administration guidelines given by the authors of the ITC-SOPI were followed.

9.4 Sample

One of this study's aims was to overcome the limitations of Study 4 (see section 8.8. in chapter 8). One of these was the lack of a specific content area to examine. A practical sampling problem arises in trying to solve this problem – what population has varying degrees of knowledge of one well-defined content area, for which there exists a VE implementation that content? It was decided to use the population of computer game players who play flight simulation games. In this case, the content area is well defined (aviation), and knowledge of it can be reasonably measured – by either asking subjects to report on their level of knowledge, or by examining their interest in other activities related to the content area (reading aviation books, visiting aviation web-sites, etc.). One can also determine if subjects have generalized or specific knowledge with relative ease (see measures in section 6.6 below). Finally, there are a number of high-fidelity flight simulation games which implement various aspects of aviation.

A total of 503 responses were collected (see section 9.3 above for a description of the recruitment procedure). Of these, 461 (91.6%) reported using Microsoft Flight Simulator 2004: A Century of Flight during their last simulator session. These were selected as the sample. This was done as a control for software platform across users. The sample, although voluntary, consisted of only men. This gender distribution probably correctly represents this particular population; as reported in section 8.3 (chapter 8), the 2003 edition of the AVSIM.com yearly user survey found only 2.6% of a sample of 14,247 were women (avsim.com, 2003). It was therefore decided to exclude gender as a variable from this study. This decision has some justification - a recent review which examined nine studies considering gender in presence (Youngblut, 2003) found a difference in only one of those studies. In terms of age, the sample was far more diverse; the mean age was 31.7 years, with an impressive range of 12 to 65 (S = 13.07 years).

9.5 Measures

Presence was measured using the ITC Sense of Presence Inventory (ITC-SOPI - Lessiter et al., 2001). The choice of this measure is discussed in section 2.6 in chapter 2. The ITC-SOPI scores were predicted by ten content knowledge, cognitive and general control factors which were defined for this study (see Table 9.1 below for a summary, the full questionnaire in appendix D). The factors are:

Thematic inertia: The same concept used in Study 4 (see section 8.1.3 in chapter 8). It is the tendency for a subject to engage in thematically related activities (e.g. reading about aviation, as well as playing aviation related games). It measures the rate at which activation decays in knowledge clusters.

Priming: The degree to which subjects engage in behaviors before playing to set a cognitive context for the simulation experience (Nunez & Blake, 2003d). This includes reading aviation books, manuals, aeronautical charts, or similar immediately before a simulation session. It is hypothesized that priming and thematic inertia are closely related: thematic inertia is a tendency or cognitive style (associated with the extent of knowledge structures), while priming a behavior which expresses that tendency.

Content knowledge: Specific knowledge of the real places and aircraft being simulated in the game. This factor is possible to implement in this study, as the game chosen for investigation provides a virtual version of the entire world and a vast number of aircraft, allowing all subjects to have real-world familiarity with the content of the game. This factor allows an estimation of the effect on presence of the fit between the simulation display and specific content expectations (see prediction 1 in section 9.2 above).). It should be noted that the items used in this factor do not ask direct, content questions (e.g. "which of the following is the oldest airplane flying today?"). A content based measure of this form could by necessity only ask very general questions, because a very specific question could still be answered incorrectly by an expert but of a different sub-field. But, if one has only very general questions, it is likely that even a non-expert could obtain a high score. Therefore, this measure looks at *exposure to information sources*, as an approximation to actual degree of expertise. The assumption is that high levels of exposure to information sources may be associated with expertise, and low levels of exposure to information sources will be associated with low levels of expertise.

Hobby cluster: The degree to which subjects engage in other aviation related activities, such as building model aircraft or reading aviation publications. This represents general knowledge of simulated content, in contrast to the content knowledge factor, which measures knowledge of specific content.

Simulator mechanics knowledge: The subject's knowledge of how simulation software works. This is measured directly by asking subjects to estimate their own knowledge, as done by Lessiter et al. (2001); in addition, it was measured by asking subjects to report on the number of modifications or simulation content they have created, based on the assumption that being able to create simulator content requires knowledge of how the simulator works. This factor can be used to control for information relevance, as it includes knowledge of the medium as opposed to knowledge of the content.

Presence management: This factor measures the presence management strategies defined in Study 4 (see sections 8.5 and 8.7.3 in chapter 8). These are steps taken by the subject to improve the immersion of the hardware platform, and their efforts to reduce distracters. The factor was expanded to include the use of consumer grade simulation input devices (joysticks, control yokes, rudder pedals, throttle quadrants, etc.). These devices not only provide improved control for the user, but also act as passive haptic devices (Insko, 2001), by mimicking the shape of real aircraft controls.

Evaluation of simulator realism: This is a measure of how realistic the subject considers flight simulations to be, in general terms. Notice that we do not use this as a measure of the realism of the system, but of the perceived realism. This cognitive factor represents arguably the most abstract level of expectation. Subjects who rate a simulation as realistic are presenting a positive interpretation bias; it can therefore be inferred that subjects who score high on this factor are less likely to interpret simulation artifacts as detracting from the experience.

Enjoyment: How enjoyable the subject finds simulations in general. This is an important control, as there is evidence to suggest that presence varies with enjoyment of the experience (Bystrom & Barfield, 1996; Nichols et al., 2000). Given that the subjects use flight simulations voluntarily for recreation, it is likely to be a factor. It is also possible that subjects who find the experience enjoyable would have a bias to overestimate their presence (the converse bias is also technically possible, but due to the self-selection of this sample, it is unlikely).

Experience-measure delay: As this study takes the unusual step of asking subjects to complete the ITC-SOPI with regard to their last flight simulation experience rather than immediately after the experience (see section 9.3 above for the procedure), this factor controls for possible memory or delay effects. The granularity of this measure was chosen as one day.

Age: This controls two factors: the possible natural covariance of age with cognitive factors (such as attention, spatial ability, etc.), as well as for the a possible correlation between age and presence (as reported by Barfield & Weghorst, 1993; Youngblut & Perrin, 2002).

9.6 Analysis strategy

The usual survey research analysis involves creating a single model from a set of predictors. This is then evaluated by examining its fit to the data (as was done in Study 4). The analysis in this study goes one step further by comparing the fit of *two* models: The first model is based on the CLCC model, and includes three sets of factors (Table 9.1): display and attention factors (form related); content and cognitive factors (content related) and general control factors. The second model is a reduced, conservative model (Table 9.2) including only display and attention factors (form related) and the general control factors. This comparison allows an evaluation in terms of the content-form debate: in essence, the model including content and cognitive factors represents the content position of the argument, while the conservative model represents the form position. Statistically, such a comparison is simple – one can perform a significance test on the difference between the error variances of two models (Neter *et al.*, 1988), which in effect computes an inference around the model R^2 values.

9.7 Results

The data were analyzed using multiple regression analyses for each of the four ITC-SOPI factors. Each of the four ITC-SOPI factors were modeled using the ten-factor model (see Table 9.1 below), and the four-factor conservative model (see Table 9.2 below).

Factor (Number of items)	Factor type	alpha	Sample item
Thematic inertia (5)	Cog	0.79	Reading about real world aviation or flight in a book, magazine or web-page makes me want to play a flight simulator.
Priming (5)	Cog	0.76	Before I play a flight simulator, I usually read an aviation/flight book, magazine, or web page.
Content knowledge (8)	Ks	0.63	I prefer to fly virtual flights around places which I have been to in real life.
Hobby cluster (7)	Kg	0.55	How many model aircraft have you built in the past year (scale models or radio-controlled)?
Simulator mechanics knowledge (9)	Ku	0.76	Have you ever created an aircraft (exterior model, flight model, etc.) for any flight simulator?
Presence management (8)	Control	0.69	What size of screen/display do you usually play simulators with?
Evaluation of simulator realism (6)	Kg	0.76	The experience provided by current commercial flight simulators is like the real thing.
Enjoyment (6)	Control	0.68	I normally find playing commercial flight simulators to be a fun experience.
Experience-measure delay – (1)	Control	N/A	How many days ago was this last session?
Age (1)	Control	N/A	What is your age?

Table 9.1: Factors used to predict ITC-SOPI scores. Factor types: Control is control factor, Cog is cognitive factor, Ku is content unrelated knowledge factor, Kg is generalized knowledge factor, Ks is specific knowledge factor.

Factor (Number of items)	Cronbach's alpha	Sample item
Presence management (8)	0.69	What size of screen/display do you usually play simulators with?
Enjoyment (6)	0.68	I normally find playing commercial flight simulators to be a fun experience.
Experience-measure delay (1)	N/A	How many days ago was this last session?
Age (1)	N/A	What is your age?

Table 9.2: The four factor conservative model predictors (note this is a subset of the ten-factor model in Table 9.1)

These two models were then compared in terms of fit to more clearly show the contribution of the content and cognitive factors. When using very large samples (as in this study) statistical power is proportionately large (Neter *et al.*, 1988), and statistical significance is easily achieved as even tiny effects are detectable. This study thus uses the 0.01 level of significance rather than the usual 0.05 level.

9.7.1 Spatial factor

For this factor, the ten predictor model is significant ($F = 17.41$, $p < 0.00001$) and gives a fit of $R^2 = 0.28$. The significant predictors (at the 0.01 level) are thematic inertia, evaluation of realism, content knowledge (as a negative factor), presence management and age (see Table 9.3 below for partial correlations). The difference between the fit of this model and that of the reduced conservative model (which has $R^2 = 0.15$) is significant ($F = 12.67$, $p < 0.00001$) – see Table 9.7.

9.7.2 Engagement factor

Again, the ten predictor model is significant ($F = 30.77$, $p < 0.00001$) with $R^2 = 0.40$. The significant predictors (at the 0.01 level – see Table 9.4 below) were the same as for spatial presence: thematic inertia, evaluation of realism, content knowledge (as a negative factor) presence management and age. The difference in fit between this model and the conservative one ($R^2 = 0.26$) is again significant ($F = 17.86$, $p < 0.00001$) – see Table 9.7.

Factor	Partial correlation
Thematic inertia	0.28
Evaluation of realism	0.18
Content knowledge	-0.12
Presence management	0.19
Age	0.14

Table 9.3: Significant predictors (p < 0.01) for the spatial factor, with partial correlations.

Factor	Partial correlation
Thematic inertia	0.33
Evaluation of realism	0.13
Content knowledge	-0.14
Presence management	0.33
Age	0.20

Table 9.4: Significant predictors (p < 0.01) for the engagement factor, with partial correlations.

9.7.3 Naturalness factor

Although the ten predictor model for this factor is significant ($F = 18.14$, $p < 0.00001$) and the fit is good ($R^2 = 0.29$), the significant predictors differ from the two previous factors. As before, thematic inertia, evaluation of realism, presence management and age are significant predictors; however, content knowledge makes no contribution, and priming is a significant predictor (see Table 9.5 below). As with the other models, the difference in fit between this and the conservative model ($R^2 = 0.13$) is significant ($F = 13.65$, $p < 0.00001$) – see Table 9.7.

Factor	Partial correlation
Thematic inertia	0.24
Evaluation of realism	0.25
Presence management	0.16
Priming	0.10
Age	0.19

Table 9.5: Significant predictors ($p < 0.01$) for the naturalness factor, with partial correlations.

9.7.4 Negative effects factor

The ten predictor model is again significant ($F = 3.57$, $p < 0.00026$), as one would expect with such a large sample; however, it shows very weak fit ($R^2 = 0.07$). The pattern of predictors is also quite different – only thematic inertia and presence management are significant predictors (see Table 9.6 below). Unlike the other ITC-SOPI factors, there is no significant difference (at the 0.01 level) in model fit between the ten predictor model and the conservative model, whose fit is $R^2 = 0.04$ ($F = 2.60$, $p < 0.02$) – see Table 9.7.

Factor	Partial correlation
Thematic inertia	0.15
Presence management	0.13

Table 9.6: Significant predictors ($p < 0.01$) for the negative effects factor, with partial correlations.

9.7.5 Overall comparison of model fit

Table 9.7 below summarizes the differences in model fit (R^2) between the ten predictor model and the conservative model. The R^2 values used were adjusted to compensate for the difference in predictors in each model, as suggested by Neter *et al.* (1998). For all ITC-SOPI factors except negative effects, the differences between model fit are significant at the 0.01 level. In general, the ten predictor model explains substantially more presence variance than the conservative model (between 1.6 and 2.3 times more fit).

ITC-SOPI Factor	R^2 for ten predictor model	R^2 for conservative model
Spatial	***0.28***	***0.15***
Engagement	***0.40***	***0.26***
Naturalness	***0.29***	***0.13***
Negative effects	0.07	0.04

Table 9.7: Summary of model fits for the ten predictor and conservative models. Significant differences ($p < 0.01$) are highlighted.

9.8 Discussion

9.8.1 Support for CLCC predictions

1. *Generalized knowledge of content will increase the presence experience, while specific knowledge will reduce the presence experience* - Three types of content knowledge were measured: Specific knowledge of the content (*content knowledge*), general knowledge of the content (*hobby cluster*), and knowledge of irrelevant content (*simulator mechanics knowledge*). Specific knowledge behaved as predicted by the CLCC model – it reduced spatial presence and engagement. The effect on the spatial factor can be understood in terms of rendering fidelity expectations held by the user, as predicted in section 9.2. Given that the simulation has a set degree of fidelity, expert users should notice more mismatches between their expectations and the display, leading to a reduction in presence. More generalized forms of knowledge were predicted to produce more non-specific expectations, which is supported by the data – the *hubby cluster* factor does not affect presence scores. For the most general of all expectations (measured by *evaluation of realism* – see 9.8.4 below), the subjects experience a positive effect on their presence as predicted by the CLCC model. Irrelevant knowledge was predicted to produce no expectations, and this is indeed shown by the data – *simulator mechanics knowledge* has no effect on presence scores.

Naturalness and negative effects failed to show knowledge effects. It is simple to understand why negative effects should not be affected, as simulator sickness (a phenomenon highly similar to negative effects) is well understood to be due to a mismatch in information between the visual and vestibular systems (Kennedy *et al.*, 1988), with high level cognition playing almost no role. However, naturalness is considered under the CLCC model as one of the cognitively higher forms of presence, and should therefore be subject to content effects. One can argue that the lack of effect is partly a measurement artifact. The ITC-SOPI items measuring this factor are at an extremely non-specific level of interpretation (e.g. "the content seemed believable to me"). It is possible that for this population (flight simulation hobbyists), the lack of effect is actually a floor effect – they have enough aviation knowledge such that responses to these items vary very little. This however, is not supported by the data – the naturalness factor has a mean and variance comparable to the other ITC-SOPI factors. Another explanation is that the expectations associated with the naturalness factor are implicit rather than semantic – that is, they are expectations about how the simulation behaves and responds to input (the 'feel'), rather than being explicit expectations about the shape or layout of the physical space and its meaning (that is, naturalness may operate on the folk physics module, rather than on the declarative memory module). Although this study provides no data to support this hypothesis, it can be considered to be indirectly supported by the theoretical distinction drawn between implicit and explicit cognition, which is for instance used in implicit memory research (see for instance Graf & Ryan, 1990) and the mental models literature (for example Johnson-Laird, 1983).

2. *High level semantic processing contributes to presence at least as much as perceptual processing* – The data strongly support this prediction for all ITC-SOPI factors except negative effects (see the discussion on why negative effects is a special case in prediction 1 above). For these three factors, the ten-factor model improves the fit from explaining roughly one sixth of the variance, to explaining roughly one third. Although it is difficulty to say that the content-related factors contribute the same amount as immersion factors, it is fairly clear that content-related factors make a significant contribution as

predicted by the CLCC model. Note that although it is a general principle that regression models with more predictors generally lead to better fit (Neter *et al.*, 1988), in this case the increase in fit cannot be solely attributed to this phenomenon, as the differences in fit are subject to a hypothesis test at the 0.01 level; however this phenomenon does make it difficult to estimate the exact contribution of the content factors. Validation of the technique used in this prediction comes from the lack of fit improvement in the negative effects factor: negative effects is hypothesized to have no semantic contribution, and indeed shows no improvement under this analysis method.

9.8.2 Implications of the design and sample

9.8.2.1 Possible effects of experience-measurement delay

The survey format of this study required violating the requirement that the ITC-SOPI be administered immediately after the experience. Although there exists no theoretical explanation as to why an experience-measurement delay should introduce systematic error into the presence measure, this study controlled for possible effects by measuring the delay. Analysis revealed the experience-measurement delay was not a significant predictor of any of the ITC-SOPI factors. One may conclude that with a quantification granularity of one day, no experience-measure delay effects exist.

9.8.2.2 Positive bias due to enjoyment

As most of the participants in this study were likely to be habitual simulation users, it was necessary to control for a possible positive response bias. The reported degree of positive bias (*enjoyment* factor) was, as expected, high – a mean score of 29.7 (on a scale ranging from 6 to 42), but was not significantly skewed. It is important to note that enjoyment was not a significant predictor of any of the four ITC-SOPI factors. One can therefore conclude that the reports of presence given by our sample were not unduly affected by their enthusiasm for the content.

9.8.2.3 Gender and age of participants

Central to any modeling study is a large representative sample (Rosenthal & Rosnow, 1991). In terms of age, this is an extremely good sample; it provides a wider age range than any comparable study in the literature (compare for instance with Study 4, Barfield & Weghorst, 1993; Youngblut & Perrin, 2002 which use large samples, but composed of university students). The same cannot be said of the gender composition of the sample. This population (flight simulation users) was chosen as it is a group of habitual VE users who have knowledge at several levels of relevance of the content of the VE. It is unfortunate that a population which is so useful for this purpose has such a significant gender imbalance, and raises the issue of whether these findings can be generalized to women. From a cognitive perspective, the most serious concern is that of possible gender differences in spatial abilities, which would have an impact on presence (Wirth *et al.*, 2007). Currently, there is no definitive evidence for the existence of such a gender difference. Some time ago the evidence seemd to suggest a difference which was diminishing over time (Stumpf & Klieme, 1989). Later meta-analyses revealed a more inconclusive picture (Coluccia & Louse, 2004; Voyer *et al.*, 1995). It is therefore possible that some of the findings of this study may not generalize across the genders.

Factor	**Spatial**	**Engagement**	**Naturalness**	**Negative effects**
Thematic Inertia	0.28	0.33	0.24	0.15
Evaluation of realism	0.18	0.13	0.25	-
Content knowledge	-0.12	-0.14	-	-
Presence management	0.19	0.33	0.16	0.13
Priming	-	-	0.10	-
Age	0.14	0.20	0.19	-
Overall R^2	0.28	0.40	0.29	0.07

Table 9.8: Comparison of partial correlations and model fit for the significant predictors of the ten predictor model on ITC-SOPI factors (empty cells indicate the predictor was not significant at the 0.01 level)

However, many of the theoretically important findings of this study (such as the contributions of content knowledge), probably rely more on semantic processes than spatial ones (as argued in chapter 4), so even if gender differences in spatial abilities exist, they may not negate the described effects.

9.8.2.4 Degree of control over display variables

This study attempts to separate the influence of form and content factors by measuring presence management practices, as well as sampling only those who use a single software platform. It should be noted that although the software was kept constant, it is still possible to have slight variations in terms of content as well as display. The software selected (as with all computer games) allows users to trade display fidelity for simulation update rate by adjusting display parameters within a narrow range. The exact degree of visual fidelity which any particular subject experienced during their last simulator session is therefore unknown. However, the range of such modifications allowed by the software is small, so no major variations in fidelity could arise form this source. Also, it is important to recognize that although this can be correctly understood as a threat to the internal validity of this design, the use of a large sample of subjects reporting on their experiences with their usual gaming situation gives this study an enviable degree of external validity (Rosenthal & Rosnow, 1991).

9.8.3 Thematic inertia and priming

The findings around the role of thematic inertia are largely consistent with the findings of Study 4. Thematic inertia is a significant predictor of all four ITC-SOPI factors, and in all cases its contribution to the ITC-SOPI factor is either higher or only slightly lower than that of presence management. The presence management factor was includes measures of display size and passive haptics, which have been established as important factors in presence (Hendrix & Barfield, 1996a; IJsselsteijn *et al.*, 2000; Insko, 2001); however, the current data suggest that if one holds the software platform constant, then thematic inertia is on average at least as important as these factors. One explanation for this phenomenon is that high thematic inertia leads to more benefit from better displays, and thus leads to learning presence management strategies. This explanation is however unlikely, as the overall correlation between presence management and thematic inertia, although significant, is fairly low ($r = 0.31$; only 9.6% of the variance is shared). Also, if this were the case, we would not expect to see both of these factors appear as significant predictors in the multiple regression, as they would share a high degree of variance. An explanation which is consistent both with the data and the CLCC model is that while presence management can be learnt, thematic inertia is probably a feature of declarative memory and therefore inherent. As discussed in section 9.2 above, thematic inertia likely contributes to presence through enabling the spreading of semantic activation; it is therefore relatively independent of perceptual factors which are associated with processing the display. Subjects who are fortunate enough to have high thematic inertia *and* engage in presence management strategies would undoubtedly have the highest presence scores, as they would achieve activation of the temporary structures in working memory from both bottom-down and top-down sources.

Intriguing is the lack of a priming effect, given that priming has been found to affect presence when manipulated experimentally (Nunez & Blake, 2003a). Priming seems only to affect naturalness, and then only slightly so (partial correlation of 0.1). One explanation is that priming is effective, but subjects do not make use of it. However, this is at odds with the evolution of presence maximization strategies argued in (and empirically supported by) Study 4. Closer examination of the data reveals that re-computing the regressions after removing thematic inertia as a predictor, leads to priming becoming a significant predictor of all ITC-SOPI factors except negative effects. This suggests that priming shares more variance with thematic inertia than with the ITC-SOPI factors. It seems reasonable to suggest that thematic inertia measures an automatic quality of cognition, where exposure to one stimulus in a content area (a book on aviation) automatically activates cognitions about related stimuli (a flight simulation) probably by means of spreading of semantic activation. On the other hand, priming measures active engagement in behaviors. It seems reasonable that without the tendency measured by thematic inertia, priming would not be effective. Thus, subjects without the tendency would not have evolved the behaviour. Also, not all those who have a high degree of thematic inertia would necessarily engage in priming behaviors for any number of practical reasons (limited time, lack of priming materials, etc.). Priming would therefore have a much higher degree of error variance than thematic inertia. We can therefore expect thematic inertia to be a better predictor of presence in the regression analysis than priming.

9.8.4 Perception of realism as an expectation

How display realism and simulation fidelity contribute to presence has been discussed in the literature to a satisfactory degree (Lessiter *et al.*, 2001; Meehan *et al.*, 2002; Slater & Steed, 2000). This study however examined the degree of realism *perceived* by the subject. As argued by Slater and Wilbur (1997), the fidelity or realism of a scene could be measured completely objectively by describing the various display parameters. For content, one could similarly measure fidelity in terms of variations between the simulated model and the actual phenomena being simulated. However, whether a subject perceives a simulation as realistic is a different matter; it depends on their expectations of the content. Asking a subject for an assessment of simulation realism is a measure of a very general, high level expectation of the experience the simulation will deliver. In this study, the simulation platform was kept constant across participants, and hardware differences were measured and controlled for. It is reasonable to argue that any differences in perceptions of realism can be attributed to cognitive factors. Also, the e*valuation of simulator realism* items asked not about the perceived realism of the last simulator session (which would have been an actual measurement of the system), but of simulation software in general. One can assume that such an average evaluation of realism would have existed before the subjects played their last simulator session, and therefore, it would have acted as an expectation during that session. As this expectation is extremely generalized, it is should be possible to satisfy even with a desktop simulation. This is indeed borne out by the data – the evaluation of realism is a positive predictor (with high partial correlations) of all factors of the ITC-SOPI except negative effects.

9.8.5 Contribution of content-irrelevant knowledge

A final interesting effect which appears in this data (or rather, fails to appear) involves content irrelevant knowledge (measured by *simulator mechanics knowledge*). Even though this study used a large sample which would typically reveal even the smallest effects, no effects were found. One can conclude with confidence that irrelevant information neither contributes to nor detracts from presence. This, together with the contribution of relevant information presented in section 9.8.1 above, supports the CLCC prediction that presence is constructed from a select subset of perceptual and conceptual information which is semantically related. It also supports to a lesser degree the notion that processing resources are allocated selectively in terms of that construction, effectively expressing a bias for content relevant information, while excluding irrelevant information. This suggests that presence is not due to a ‘willing suspension of disbelief’ (discussed in Botella *et al.*, 2003; Ryan, 1999), but rather that information which might lead to disbelief is filtered out, so that the only aspect of willingness in the presence experience may the decision to engage with the VR system.

9.9 Conclusion

9.9.1 Interaction of expectations and mediated content

As discussed in 9.8.4 above, a high evaluation of realism was associated with an increase in presence, due to the general nature of the expectation associated with that evaluation. Conversely, for content knowledge, a negative relationship existed with presence, due to the highly specific expectations associated with specific content knowledge. At the same time, knowledge not associated with the simulation content (such as knowledge of the simulation mechanics), had no effect, as it creates no expectation for the content.

The CLCC model explains how these expectations operate in presence. Before entering the experience, the subject has expectations about the content and the experience. During the experience, the simulation provides perceptual cues on several modalities (which is all a system can do). If these cues match the expectations arising from the subject's content knowledge, then the subject will have a coherent presence experience in the system. If not, the lack of match will lead to a reduced presence experience through the attracting of attention to perceived errors in the content, as well as a reduced sense of naturalness and reduced engagement with the material. Such a mismatch is more likely to occur in the face of highly specific expectations, and less likely in the face of generalized expectations. This leads to a general expectation principle: VE relevant knowledge creates a cognitive context of expectations, with more knowledge leading to more specific expectations; and presence is more likely to occur when expectations are matched by the VE system. This expectation principle is well supported by this data. If one orders the content relevant factors with respect to how specific an expectation they generate and plots them against their partial correlations (Figure 9.1 below), then the predicted downward trend is discernible for all the ITC-SOPI factors except negative effects.

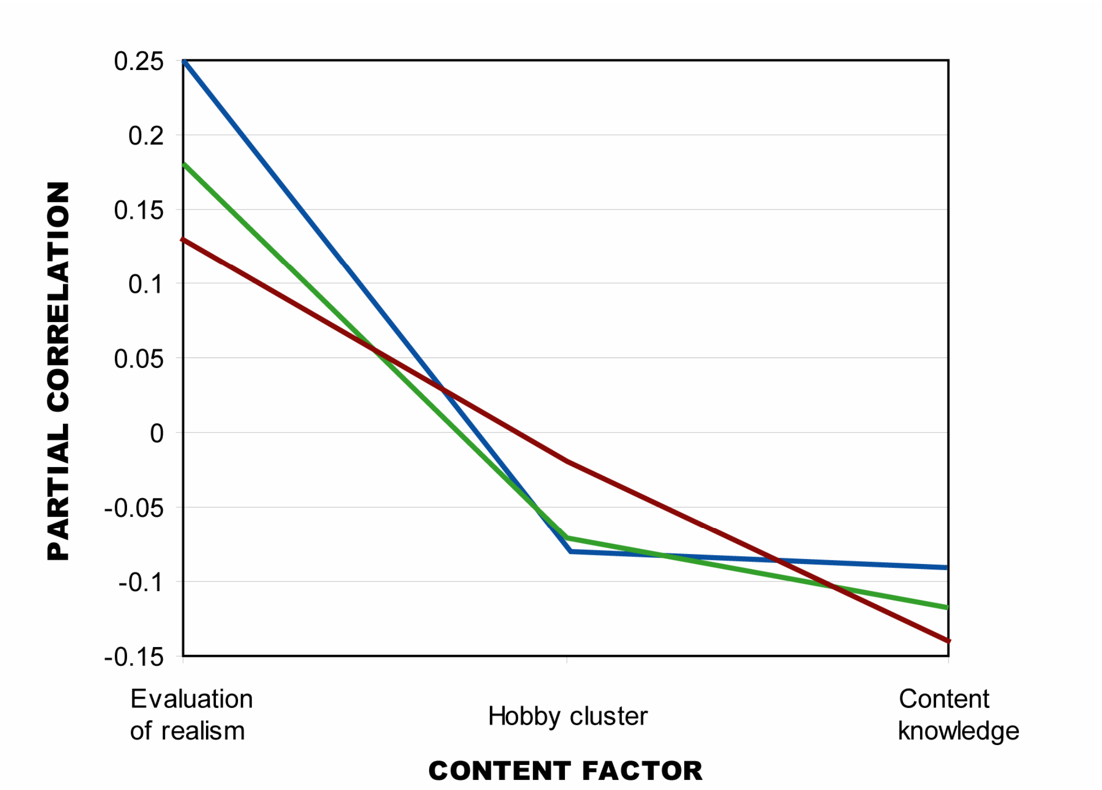


Figure 9.1: Content factors, in decreasing level of generality of expectation plotted against their partial correlation with three ITC-SOPI factors: Spatial is the blue curve; Engagement is the red curve; Naturalness is the green curve.

9.9.2 Relative contribution of content and cognitive factors

In general terms, the data show that the content and cognitive factors add significant fit over the conservative model (see Table 9.7 above). The conservative model replicates the large body of published work which argues for the importance of display related factors in presence (such as Barfield & Hendrix, 1995; Insko, 2001; Sas & O'Hare, 2001 and others; Slater & Wilbur, 1997). However, the difference in fit between the models highlights the importance of considering content related factors when predicting presence. It is unlikely that the significant increase in model fit brought about by adding the content related factors is simply an artifact of the selected display factors, the addition of more predictors to the regression, or of the method used.

In general terms, the data show that the content and cognitive factors add significant fit over the conservative model (see Table 9.7 above). The conservative model replicates the large body of published work which argues for the importance of

display related factors in presence (such as Barfield & Hendrix, 1995; Insko, 2001; Sas & O'Hare, 2001 and others; Slater & Wilbur, 1997). However, the difference in fit between the models highlights the importance of considering content related factors when predicting presence. It is unlikely that the significant increase in model fit brought about by adding the content related factors is simply an artifact of the selected display factors, the addition of more predictors to the regression, or of the method used.

Chapter 10: Study 6 – Effects of semantic content manipulation on presence

Although Studies 4 and 5 provided evidence for the roles of top-down information and content effects in presence, the use of surveys prevented inferences of causality. Study 6 overcomes this limitation by manipulating semantic information, to test the causality of context and content expectation on presence.

This study examines the interaction of content from three sources: Priming which occurs before the VE experience, VE content (which remains constant across subjects), and non-diegetic music which plays during the experience. VE content was not manipulated because it is unknown on which dimensions it varies. It would have been possible to adopt the strategy used in Study 5, and select only subjects which have an interest in or knowledge of the content chosen – but that technique precludes random sampling (which is essential to an experimental design – Rosenthal & Rosnow, 1991). It was therefore decided to use a single VE, with true randomization of subjects across conditions. Although the content of the VE is constant, it still plays an important role in the presence experience, as it provides the semantic context against which the priming and non-diegetic music manipulations can be evaluated.

10.1 Semantic priming

Priming is a process which prepares subjects for some cognitive task (Maxfield, 1997), and is generally used to determine if some cognitive process has a top-down component. The reasoning behind this method is that if a task involves top-down processing, then providing a subject with relevant semantic information will pre-activate top-level structures and enhance processing; conversely, if a task does not make use of top-down processing, then priming will have no effect on that task (Rumelhart *et al.*, 1986; Tulving & Schachter, 1990). Some work has been done on priming in presence. Nunez & Blake (2003a) found an interaction effect between semantic primingand display fidelity on SUS and PQ scores, suggesting that presence may have a top-down component. The current study uses the same manipulation.

10.2 Non-diegetic music

Non-diegetic music is a film-making term which describes music which does not arise from the space in which the action occurs (Monaco, 2000). Although not located spatially in the film environment, it is usually semantically or emotionally related to the events portrayed on the screen. Non-diegetic music is interesting for presence because it reduces the fidelity of the display, but (as Study 4 suggests) seems to be a contributor to presence, at least for veteran gamers. What effect non-diegetic music will have on presence is not clear from the models. The three-pole model (Biocca, 2003) would predict that non-diegetic music would reduce presence, as it would pull the presence away from the VE pole towards the mental imagery pole. Similarly, the environment selection model (Slater, 2002) would predict that non-diegetic music would reduce presence as it would be a non-immersive distracter. On the other hand, the FLS model (Waterworth & Waterworth, 2001) would predict that non-diegetic music could stimulate the sensus dimension, enhancing presence; similarly, the MEC model (Wirth *et al.*, 2007) would predict that non-diegetic music would help direct focused attention on the VE, and the music might interact with domain-specific interest to enhance presence.

10.3 Predictions about semantic content made by the CLCC model

1. *Priming (with material semantically related to the VE content) will enhance presence, particularly the cognitively higher forms (such as engagement and naturalness)* – During priming, knowledge clusters activate as the subject engages with the priming materials. When priming ends, the knowledge clusters remain active for a time due to thematic inertia, such that when the subject experiences the VE, the knowledge structures will still be partly activate. In the CLCC model, presence relies on the semantic bias of the system which is fed down from active knowledge structures. Normally, the semantic bias comes about as temporary structures in working memory activate knowledge clusters in declarative memory. If the knowledge clusters are already partly activated by priming, this should lead to a more substantive bias, as the knowledge clusters receive activation from two sources (priming and temporary structures) rather than just one. The enhanced bias should lead to enhanced presence. This effect will be magnified for cognitively higher

forms of presence because of the larger role played by expectation and semantic processing in those forms.

2. *Non-diegetic music which is semantically related to the VE content will enhance presence (particularly the cognitively higher forms such as engagement and naturalness)* – The CLCC model does not differentiate between stimuli originating from the VE or from other sources (as origin is inferred during the construction process). This means that non-diegetic music (which can be considered as reducing the fidelity of the system) can still contribute positively to presence, provided it is semantically related to the current model bias. Such music will be included in the model bias, effectively increasing it, and thus enhancing presence (particularly the cognitively higher forms, which rely more on semantic processing).

3. *Non-diegetic music which is semantically unrelated to the VE content will be filtered out and have no effect on presence* – This prediction is related to prediction 2 above. All data are evaluated relative to the semantic bias, so semantically unrelated data will either be filtered out at the stimulus attenuator (which acts to maintain the semantic bias), or will lead to a reconstruction which includes the new data. If the music has a relatively constant volume and intensity, it is unlikely to force its way through the stimulus filter, so the prediction is that semantically unrelated music will be filtered out, and therefore have no effect on presence.

10.4 Sample

181 undergraduate students participated for course credit (145 women and 36 men). Mean age was 21.45 years (S = 3.46). The sample was measured on computer experience, game playing experience, and knowledge of virtual reality (using a 6 point scale, 0 = no knowledge/experience, 5 = expert). The results indicate the sample were novices (see Table 10.1 below).

	Mean	SD
Computer experience	1.629	0.597
Game playing experience	0.204	0.800
Game playing frequency	0.756	0.742
Knowledge of VR	0.326	0.585

Table 10.1: Subjects' knowledge and experience of VR related technology

10.5 Apparatus

Five subjects were run simultaneously in a dedicated space. The desktop machines used (see Table 10.2 below) produced a measured update rate ranging between 17Hz and 28Hz.

Hardware	
Display:	17" Samsung Syncmaster 750 CRT
Graphics card:	GeForce 6200, 128MB RAM
Processor:	Intel Pentium 4, 2.8GHz
RAM:	512MB, DDR333
Input devices:	Keyboard and optical mouse
Sound:	Stereo, by headphone

Table 10.2: Hardware specification of the desktop machines used

10.5.1 Virtual environments

The study used two VEs - a training VE (the same used in Study 1 – see 5.4.1 in chapter 5, but with a different task) and the main VE in which the study was conducted. The main VE represents a European monastery, and contained two buildings (the monastery and a chapel) – see Figure 10.1 and Figure 10.2 below. The VEs were rendered using the Genesis3D engine (http://www.genesis3d.com), at a resolution of 1024x768x32. Audio was spatialized (through headphones), and the subject could hear their footsteps in the VE as they walked. Control in the VE was by the Quake Keys method (Dalgarno & Scott, 2000).

Figure 10.1: Image from the monastery environment (upstairs landing)

10.6 Procedure

When subjects arrived, the researcher explained that the study was examining psychological aspects of virtual reality. Subjects were then randomly assigned to a condition (see 10.8 below for the manipulations). They completed an introductory questionnaire measuring biographical information and their level of VR experience. They were then shown the controls and task (described in 10.7 below), and taken through the training VE. The subjects were then, according to the experimental condition, given priming materials to read on-screen (see 10.8 below for a description of the materials).

The subjects then read brief scenario for their experience – they were told that they were assistants to an anthropology professor, who had negotiated with the monks of a local monastery for their collection of rare books. The subjects' task was to go to the monastery and search for and collect the books (see appendix E for scenario text). The main VE was then started, and subjects performed the task for a timed period of 15 minutes. The subjects were administered the ITC-SOPI, followed by the Izard emotion scale (Izard, 1991). See measures (10.9) below.

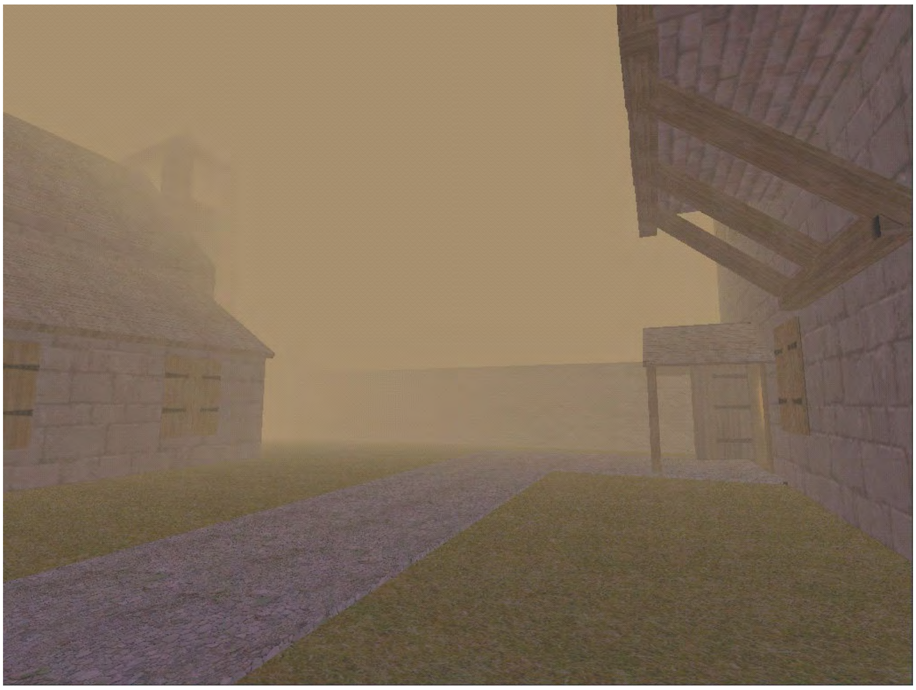


Figure 10.2: Image from the monastery environment (courtyard)

10.7 Experimental task

The task was object search and collection, and was semantically concordant with the VE content and scenario. A total of 30 objects (ancient books) were placed in the VE, which the subjects had to search for and retrieve. Subjects retrieved an object by moving close to it, centering it on screen (see Figure 10.3 below), and clicking the mouse button. At any time the subject could get a count of how many objects were left to retrieve. However, this object counter never displayed a number less than 1; this prevented subjects who completed the task faster than others from being idle until the end of the 15 minute session.

10.8 Content manipulations

This study manipulated two factors: Priming (Nunez & Blake, 2003a) and non-diegetic music fit. Priming was manipulated by giving subjects a set of materials to read before the VE experience (as used in Nunez & Blake, 2003a). Two priming conditions were used: VE-relevant (see Figure 10.6 and appendix F) and VE-irrelevant (see Figure 10.5 and appendix F). The materials are taken from Nunez & Blake (2003a), as these have been shown to produce differences in SUS and PQ scores. They have also been shown to take equivalent times to read by undergraduate samples (Nunez, 2002).

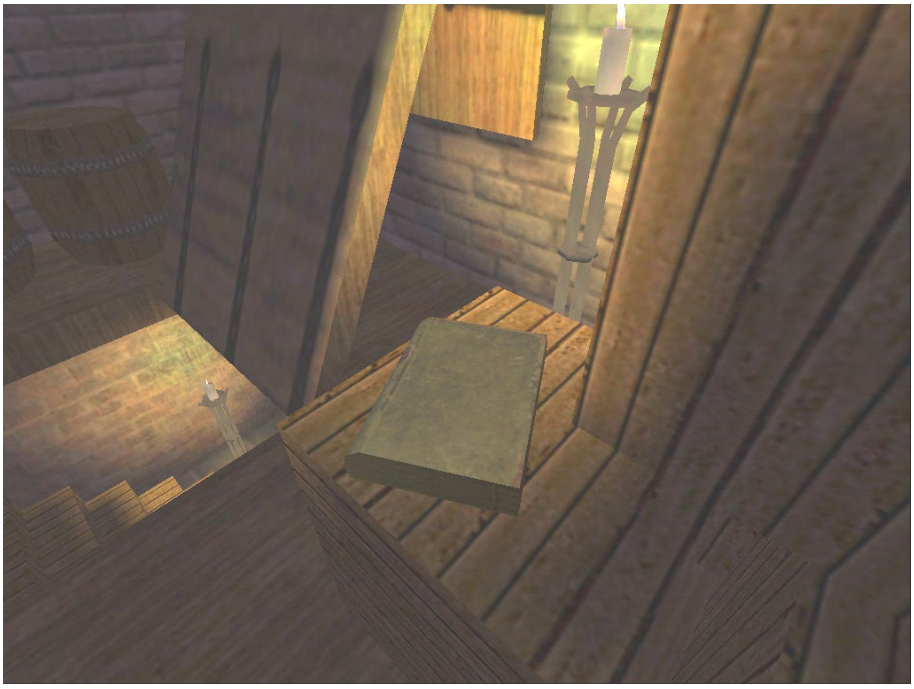


Figure 10.3: A book centered on the screen about to be collected (with a mouse click)

Non-diegetic music fit was manipulated into three conditions: *fit, no-fit* and *baseline*. In the fit and no-fit conditions, the subjects heard background music during their entire VE experience (a four minute piece looped to fill the entire 15 minute session). In the baseline condition, no music was played. To select the music, a set of seven pieces which were though to be of high or low fit to the VE were first selected. These were then presented to ten independent judges together with a set of fifteen screenshots of the monastery VE.

The judges were asked to rate, on a seven point scale, the match between the theme of the images and of the music (1 = do not fit together at all, 7 = fit together perfectly). The pieces with the highest and lowest average ratings were then chosen for the fit and no-fit conditions. The best fitting music (mean fit 6.0) was *Salvator mundi, salva nos*, a medieval choral piece by Thomas Tallis; the worst fitting piece (mean fit 1.7) was *Nin wa itsumodemo issho,* an upbeat recorder piece from a children's television program, by Masaki Kurihara.

Figure 10.4: Sample of the VE-irrelevant priming material (trains; Nunez & Blake, 2003)

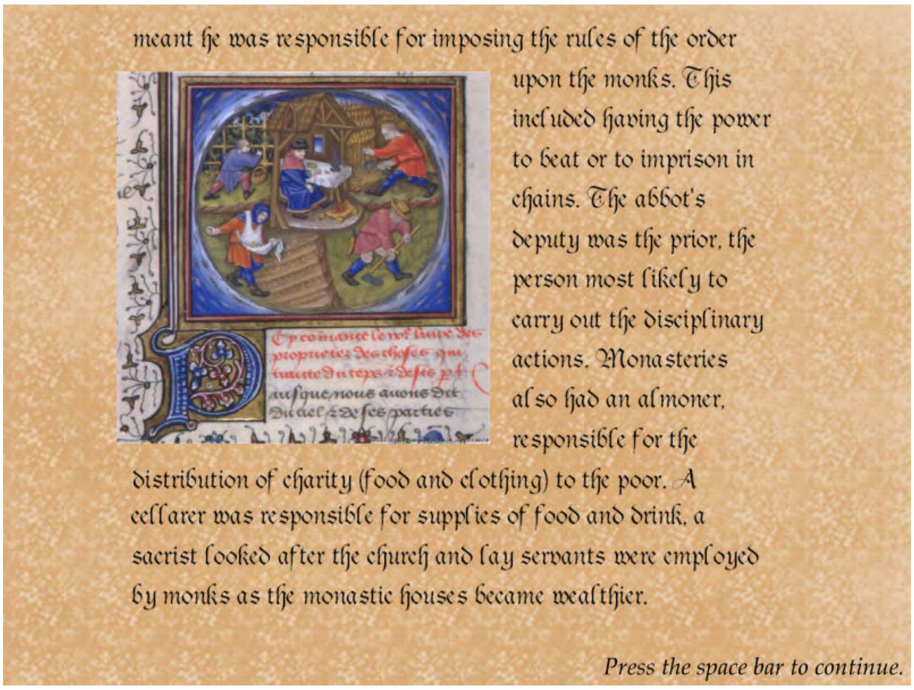


Figure 10.5: Sample of the VE-relevant priming material (monasteries; Nunez & Blake, 2003)

10.9 Measures

Presence was measured using the ITC-SOPI (Lessiter *et al.*, 2001; see also 2.6 in chapter 2). A potential confound in this study is emotion. Bever (1988) argues that music, by its tone and rhythm, can encode emotions which are widely recognized; this is echoed by Pinker (2002), although with the qualification that this effect is culture bound. A number of empirical studies support this idea (Sloboda, 1991), with particularly mode, rhythm and tempo playing an important role in producing particular emotions (Gabreilsson & Lindström, 2001). Because emotion is theorized to be a significant factor in, this is a concern: Alcañiz *et al*. (2003) and Riva *et al*. (2003) have proposed that emotion acts as mediator to VE content during presence. Some empirical support exists for this relationship. Baños *et al*. (2004; , 2004) showed that presence correlates with emotion, and that presence scores can be changed by manipulating the emotional tone of the environment.

Because this study manipulates music, it is possible that the music selected will lead to differences in presence due to emotion effects rather than semantic effects as intended. This was a particularly concern as the music chosen for the fit condition was dark and moody, while that chosen for the no-fit condition was light and upbeat. As a control for this, the study included a measure of emotion response - the second edition of Izard's differential emotions scale (DES-II Izard, 1991), which consists of 30 items, each either a word or phrase describing an emotion. Subjects rate the degree to which they felt that emotion on a seven point scale (1 = Not at all, 7 = Very much). The DES-II has been recently validated for research use (Fuenzalida *et al.*, 2005). The items were further collapsed into three factors: Positive and Negative emotions (following Izard, 1991), and attention focusing – see table Table 10.4 below for which items compose these factors, and 10.10.2 for a description of the procedure.

10.10 Analysis & results

10.10.1 Task performance

The task was designed to keep attention focused on the VE, so task performance is not central to the validity of the study. The task was in fact very easy – 139 subjects (86.75%) collected 25 or more of the 30 objects (see Figure 10.6). Due to the ease of the task, 42 subjects who collected fewer than 25 objects were excluded form analysis, as their poor task performance may indicate problems with the interface or other system failures rather than natural variation in task performance. A chi-square analysis shows that the subjects excluded were evenly distributed in the design ($\chi^2 = 1.400$, $p < 0.496$). The exclusion will therefore not bias subsequent analysis.

The distribution of subjects included in the design is shown in Table 10.3 below; the distribution is even across cells ($\chi^2 = 0.548$, $p < 0.760$). Although 77.5% of subjects were women, the men are distributed evenly in the design ($\chi^2 = 0.881$, $p < 0.643$).

Priming	***Non-diegetic fit***		
	Fit	**Baseline**	**No fit**
VE-relevant	28	27	23
VE-irrelevant	25	30	27

Table 10.3: Cell frequencies (N) for each of the six conditions.

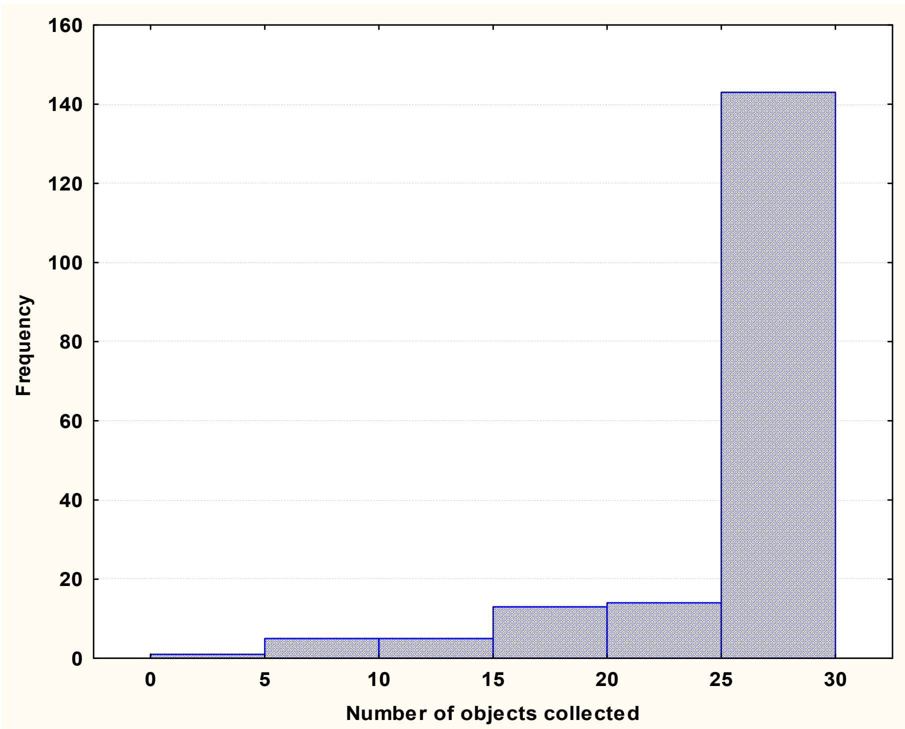


Figure 10.6: Histogram of task performance (number of objects collected)

10.10.2 Factorization of Izard emotion scale

The DES-II scale is theorized to contain two factors - positive and negative emotions (Izard, 1991). The DES-II data from this study was factor analyzed to extract two factors (using varimax rotation). The factor loadings on the two resulting factors were taken by considering items which had a factor loading higher than 0.7. The resulting factor structure (shown in Table 10.4 below) supports the theoretical distinction between positive and negative emotion. The negative emotion factor explains 0.23 of the total variance (eigenvalue 6.983), while the positive factor explains 0.20 of the total variance (eigenvalue 5.740). A third factor was constructed *a-priori* from the DES-II, to measure attention focus.

Positive emotion ($R^2 = 0.23$)	**Negative emotion** ($R^2 = 0.20$)	**Attention Focus** (*a priori*)
Happy	Disgusted	Alert
Joyful	Disdainful	You were concentrating
Surprise	Downhearted	Attentive
Amazed	Angry	
	A feeling of revulsion	
	Scornful	

Table 10.4: DES-II items included in each of the factors. The positive and negative emotion factors were produced by a factor analysis; the attention focus (not a standard DES-II factor) factor was produced a-priori.

10.10.3 Manipulation effects on emotion

Each of the three DES-II factors were analyzed using a 2x3 (priming x non-diegetic music fit) factorial analysis of variance

10.10.3.1 Positive emotion

There was no significant main effect of priming condition ($F(1, 175) = 0.0136$, $p < 0.907$) or of non-diegetic music fit condition on positive emotion ($F(2, 175) = 0.195$, $p < 0.822$). There was also no interaction effect between these factors on positive emotion ($F(2, 175) = 1.0036$, $p < 0.368$).

10.10.3.2 Negative emotion

As with positive emotion, there were no main effects (priming: $F(1, 175) = 0.046$, $p < 0.830$; non-diegetic music: $F(2, 175) = 0.348$, $p < 0.348$) or an interaction effect ($F(2, 175) = 0.385$, $p < 0.681$) on negative emotion.

10.10.3.3 Attention focus

Again, no main effects (priming: $F(1, 175) = 0.713$, $p < 0.399$; non-diegetic music: $F(2, 175) = 0.053$, $p < 0.948$) or interaction effect ($F(2, 175) = 0.035$, $p < 0.965$) were found on attention focus.

10.10.4 Emotions and presence

To test if emotions were associated with presence scores (as predicted by Alcañiz *et al.*, 2003; R. M. Baños *et al.*, 2004), we conducted zero-order correlations between the four ITC-SOPI factors and the three DES-II factors. All correlations are significant. The results are shown in Table 10.5 below. Note that positive and negative emotion give inverted correlation directions.

Izard DES-II factor	***ITC-SOPI factor***			
	Spatial	**Engagement**	**Naturalness**	**Negative Effects**
Positive emotion	***0.56***	***0.69***	***0.47***	***-0.32***
Negative emotion	***-0.18***	***-0.26***	***-0.19***	***0.32***
Attention focus	***0.55***	***0.65***	***0.41***	***-0.31***

Table 10.5: Zero-order correlations between presence and DES-II scores. All correlations are significant at the 0.05 level.

10.10.5 Modeling presence from emotion and experimental condition

Because of the significant zero-order correlations between presence and DES-II factors, it was decided to include DES-II factors into the analysis of presence as covariates. Each ITC-SOPI factor was predicted using a general linear model (GLM) which included the priming condition, the non-diegetic music fit condition, and the three DES-II factors.

10.10.5.1 Spatial factor

The Cronbach's alpha for this factor was 0.92. For this factor, the overall model is significant: $F(8, 151) = 11.896$, $p < 0.0001$, $R^2 = 0.386$ (see Table 10.6). Analysis of the effects shows that only the DES-II positive emotion and attention focus factors are significant predictors. Positive emotion has a partial correlation of 0.327 ($R^2 = 0.387$), while attention focus has a partial correlation of 0.305 ($R^2 = 0.360$).

Effect	df	F	*p*
Izard Attention Focus	***1***	***15.577***	***0.0001***
Izard Positive	***1***	***18.117***	***0.0001***
Izard Negative	1	0.858	0.355
Priming	1	0.034	0.853
Non-diegetic music	2	0.556	0.574
Priming x Non-diegetic	2	1.673	0.191

Table 10.6: GLM results for the ITC-SOPI spatial factor. Significant effects ($p < 0.05$) are highlighted

10.10.5.2 Engagement

The Cronbach's alpha for this factor was 0.91. The overall model is significant: $F(8, 151) = 28.571$, $p < 0.0001$, $R^2 = 0.602$ (see Table 10.7). Analysis of the effects shows that only the three DES-II factors predict engagement. Positive emotion has a partial correlation of 0.515 ($R^2 = 0.387$), negative emotion has a negative partial correlation of -0.188 ($R^2 = 0.065$), and attention focus has a partial correlation of 0.406 ($R^2 = 0.360$).

10.10.5.3 Naturalness

The Cronbach's alpha for this factor was 0.79. The overall model is significant: $F(8, 151) = 7.425$, $p < 0.0001$, $R^2 = 0.282$ (see Table 10.8). This factor shows a similar pattern to the spatial factor: the positive emotion and attention focus factors of the

DES-II scale predict naturalness. Positive emotion has a partial correlation of 0.256 (R^2 = 0.387), and attention focus has a partial correlation of 0.203 (R^2 = 0.360). For this ITC-SOPI factor however, non-diegetic music fit was significant. A Fisher's LSD post-hoc test shows that there was no significant difference between the no-fit (M = 2.946) and no music (M = 2.848) conditions; but the fit condition (M = 3.240) was significantly larger than the other two conditions. As can be seen by the mean differences between conditions, the effect is small (see Figure 10.7 below).

Effect	df	F	*p*
Izard Attention Focus	***1***	***29.926***	***0.0001***
Izard Positive	***1***	***54.574***	***0.0001***
Izard Negative	***1***	***5.557***	***0.016***
Priming	1	0.046	0.829
Non-diegetic music	2	0.101	0.903
Priming x Non-diegetic	2	2.232	0.100

Table 10.7: GLM results for the ITC-SOPI engagement factor. Significant effects (p < 0.05) are highlighted

Effect	df	F	*p*
Izard Attention Focus	***1***	***6.549***	***0.011***
Izard Positive	***1***	***10.616***	***0.001***
Izard Negative	1	0.902	0.343
Priming	1	0.028	0.866
Non-diegetic music	***2***	***4.513***	***0.012***
Priming x Non-diegetic	2	0.662	0.517

Table 10.8: GLM results for the ITC-SOPI naturalness factor. Significant effects (p < 0.05) are highlighted

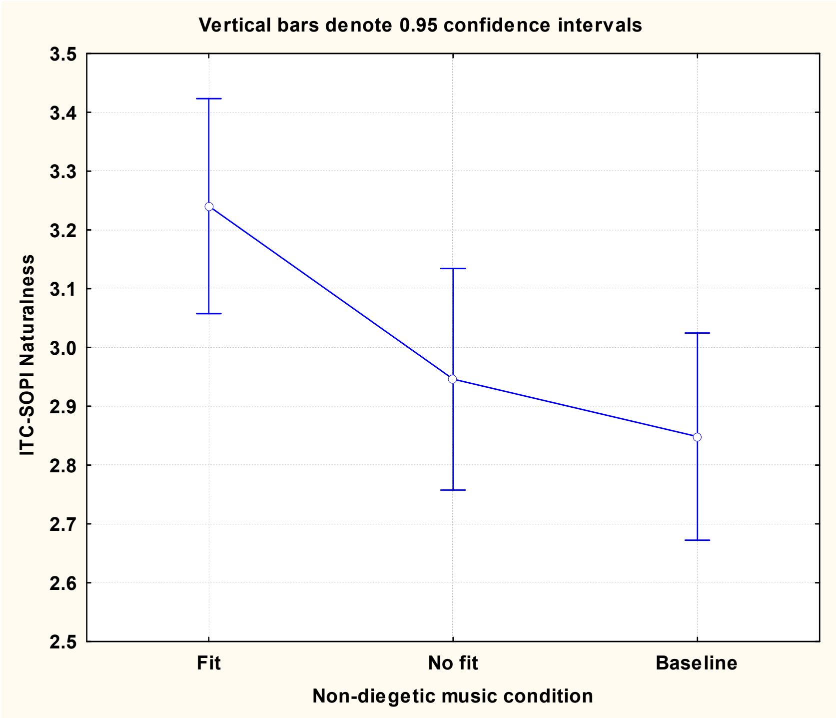


Figure 10.7: Means plot for the non-diegetic music condition effect on naturalness.

10.10.5.4 Negative effects

The Cronbach's alpha for this factor was 0.88. The overall model is again significant: $F(8, 151) = 5.220$, $p < 0.0001$, $R^2 = 0.216$ (see Table 10.9) Only the negative emotion factor and priming condition predict negative effects. Negative emotion has a partial correlation of 0.305 ($R^2 = 0.065$). Subjects primed with VE relevant material had lower negative effects scores ($M = 2.526$) than subjects primed with VE irrelevant material ($M = 2.981$).

10.11 Discussion

10.11.1 Support for CLCC predictions

In terms of the predictions made by the CLCC model in 10.3 above, the results suggest the following:

1. *Priming (with material semantically related to the VE content) will enhance presence, particularly the cognitively higher forms (such as engagement and naturalness)* – Only one of the four ITC-SOPI factors (negative effects) was sensitive to the priming manipulation (subjects primed for the VE content

experienced slightly lower negative effects scores). The negative effects factor is however problematic to explain as it relates primarily to physiological responses such as eyestrain and nausea (Lessiter *et al.*, 2001), which makes it unclear why a purely cognitive manipulation such as priming has an effect. One explanation is that the subjects who had more negative effects had more negative emotions and therefore had a poorer presence experience; but this cannot be the case because emotions were factored out of the priming effect by the GLM. It may well be that negative effects, like simulator sickness, occur due to a mismatch between expectations and incoming data (Kennedy *et al.*, 1988). This would mean that those subjects in the VE relevant priming condition would have had expectations which would more closely have matched what they experienced in the VE.

Effect	df	F	*p*
Izard Attention Focus	1	3.629	0.058
Izard Positive	1	1.583	0.210
Izard Negative	***1***	***15.556***	***0.0001***
Priming	***1***	***6.390***	***0.012***
Non-diegetic music	2	0.233	0.792
Priming x Non-diegetic	2	0.545	0.589

Table 10.9: GLM results for the ITC-SOPI negative effects factor. Significant effects ($p < 0.05$) are highlighted

Although the general lack of a priming finding contradicts Nunez & Blake (2003a), it is worth noting that that study found an interaction between fidelity and priming, and not a priming main effect. As this study did not manipulate display fidelity, there was no reason to expect a replication of that previous finding.

2. *Non-diegetic music which is semantically related to the VE content will enhance presence (particularly the cognitively higher forms such as engagement and naturalness)* – This prediction was partly supported. Although non-diegetic music did not affect the spatial and engagement factors, it was a predictor of naturalness, which is the factor which the CLCC model predicts is most closely tied to semantic expectations. As predicted, it is only the relevant music condition which

produces a benefit above the baseline (no-music) condition. This suggests that the music is activating knowledge clusters which are then receiving activation from the temporary working memory structures. Again, the benefit cannot be attributed to any emotion change produced by the music itself, firstly because there was no main effect of music on emotion, and secondly because the GLM factors out emotion. Finally, the effect cannot be ascribed to the music simply providing more information (as one would argue the multimodality effect), because it is only the relevant music which gives the benefit. One must therefore conclude that this is a content effect, and not simply an immersion or information load effect.

3. *Non-diegetic music which is semantically unrelated to the VE content will be filtered out and have no effect on presence* – This prediction is strongly related to prediction 2 above, and has a similar level of support. Again, only naturalness shows the predicted result: there was no significant difference between the baseline condition and the irrelevant music condition. This suggests that, as predicted, the semantically irrelevant music is filtered out and has no effect on subsequent processing.

10.11.2 Emotion and presence

A surprising finding was the predictive power of the DES-II. The positive emotion factor predicted three of the four presence factors (spatial, engagement and naturalness), while the negative emotion factor predicted two of the four (engagement and negative effects). Furthermore, the pattern is highly intuitive: positive emotion is positively correlated with presence, and negative emotion is negatively correlated with presence. This pattern supports Lessiter *et al*'s (2001) suggestion than engagement is associated with enjoyment of the VE (as enjoyment would no doubt generate positive emotions), but the effect seems to generalize to the other factors of presence, including the fairly automatic *spatial presence* factor. Although the link between emotion and presence has been demonstrated by Baños *et al*. (2004; , 2004), it was surprising that after factoring out all the experimental manipulations as well as attention allocation, the emotion factors still emerged as the best predictors of presence. These findings suggest that emotion is an important factor in presence, and should be included in predictive models in the future. At the very least, measures of emotion should be included into future experiments as controls for extraneous variance.

10.11.3 Fidelilty, non-diegetic information and presence

An interesting result is the lack of negative effects from adding non-diegetic music to the VE (see Figure 10.7 above for the means plot). This finding contradicts the general assumption that presence will increase as display fidelity increases (Biocca, 2003). This data shows that decreasing the fidelity of the system (by adding background music) can lead to an increase in presence, or to no change at all, depending on the semantic fit between the music and the VE content.

One conclusion which can be drawn from this study (which is supported by the CLCC model) is that presence is not simply a consequence of replacing the subject's sensory stimuli with other, similar stimuli (Slater, 1999), but rather that presence occurs when a set of sensory stimuli matches the subject's expectations for that situation. The reason non-diegetic music contributes to presence (and particularly to naturalness) is that when placed in front of a display and shown moving images, subjects have expectations associated with films and television, including an expectation for non-diegetic music which matches the images they see. This accounts for the seemingly counter-intuitive finding that subjects find the interactions more natural or realistic when fitting non-diegetic music is played – it is not more natural *when one expects* the VE to be like an unmediated experience, but it is more natural when one expects it to be like a television or film experience.

10.12 Conclusion

The data supported the CLCC predictions moderately well. Three important findings emerged from this study. First, manipulating content related variables can affect the presence experience, as suggested by the naturalness and negative effects results. It also seems that the source of stimuli (be it diegetic or non-diegetic) does not matter to the presence experience; rather, it is the semantic content of the stimuli which is important. Second, data which is semantically unrelated to the content does not necessarily interfere with the presence experience. It can be filtered out to prevent further processing, and this filtering process leads to presence scores no lower than baseline. Third, emotion is a strong correlate of presence, which should be considered both in future models and as a control in empirical studies.

Chapter 11: Conclusions

This project aimed to develop and validate the CLCC model, which is a psychologically and evolutionarily plausible cognitive model of presence. The CLCC model contains several innovations. First, it incorporates both working memory and attention which allows it to explain both breaks in presence and drifts out of presence. Incorporating working memory also allows predicting how task performance might interfere with presence, and how semantic consistency can be factor in presence. Second, it models semantic processing (by means of declarative memory and the semantic bias), which allows it to make explicit predictions about content effects. Third, the model includes media decoders as an explicit mechanism to explain both how different media (immersive and non-immersive) can bring about presence, and how literacy with media can lead to individual differences in the experience. Finally, due to the inclusion of declarative memory, the model allows for effective explanations of individual differences based on previous knowledge and expertise. Beyond these innovations, the CLCC model is capable of explaining well-established findings in the literature, such as the immersion-presence relationship, the role of attention in presence, breaks in presence, forms of presence beyond spatial, as well as all five theoretical problems (the book, physical reality, dream state, inverse presence and virtual stimuli problems). This gives it more explanatory power than existing presence models. Overall, the model exists in order to explain the top-half of the presence experience. As argued in Chapter 1, users of virtual environments are active processing agents, who bring with them their previous experiences, media literacy and other individual factors into the VE; it is therefore incorrect to consider them simply as passive elements in the presence experience (as is done by the two-pole and three-pole models of presence – see Chapter 3). Rather, one needs to consider how users process the VE in their own cognitive context. The CLCC model is able to do this by incorporating semantic memory systems, and by positing a specific mechanism (active constructions in working memory) to explain how this context interacts with immersion factors. At first glance therefore, it would seem that cognitive factors should at least be considered as important as immersion factors (and indeed, Study 5 seems to support this idea). However, one cannot ignore the abundance of empirical studies which show that immersion factors, as a trend, increase the presence experience; the interaction between immersion and cognitive factors is therefore likely to be subtle. The CLCC model, with its integration of top-down and bottom-up factors, provides a highly structured theoretical framework in which to investigate these interactions in the future.

The predictions and findings of the empirical studies conducted to evaluate the CLCC model are summarized as a table in Appendix H. The evidence from the six studies shows that overall the CLCC model is valid. The stimulus attenuator does select stimuli based on semantic as well as physical properties as predicted (see 11.1 below), and there does seem to be a general semantic bias which interacts with incoming stimuli (see 11.1 and 11.2 below). Most clear were the content effects tested – there is strong evidence that, as predicted, declarative memory plays an important role in presence, through the creation and matching of expectations based on previous experience (see 11.1 and 11.3 below). There are also some weaknesses in the model. No evidence was found for the operation of the media decoders, but working memory does seem to play a role in presence, although smaller than predicted by the model (see 11.1 below).

11.1 Working memory and attention in presence

The CLCC model is the first presence model to incorporate working memory; Studies 1 and 2 were thus conducted without methodological precedent in this field. Similarly, Study 3 was innovative in that it was the first presence study to apply the divided attention paradigm to presence. In some sense, the lack of previous work in these areas increased the risk that the CLCC model would be invalidated straight out of the box; however, the success of the working memory interference and divided attention paradigms in related fields (such as human factors – see 4.2 in chapter 4) promised not only theoretical advances, but also a potential clarification of the relationship between presence and task performance (see Bystrom *et al.*, 1999a for a description of this problem).

The CLCC model made strong predictions with regards to working memory and attention in presence. Specifically, it predicted that under all media conditions, spatial presence should use the least amount of processing resources, while engagement and naturalness should use more resources (see 5.2 in chapter 5); furthermore, when experiencing an encoded medium, all presence factors should show more working memory use (see 6.1 in chapter 6). The data support only a weak version of these predictions. In a semi-immersive VE (Study 1), spatial presence showed no working memory interference, while engagement and naturalness showed only slight interference effects; and in a non-immersive VE (Study 2), none of the factors showed any working memory interference. This lack of effect may be due to an error or artifact in the studies, because not only do they contradict CLCC predictions, but MEC model predictions also. The MEC model posits that subjects construct SSMs which are tested against incoming perceptual evidence (Wirth *et al.*, 2007). It is reasonable to assume that the construction and testing of these structures will make use of some cognitive resources, and although the MEC authors do not state what resources are involved in this task, working memory seems a likely candidate. Therefore, the lack of working memory effects may not point to a weakness in the CLCC model, but rather a weakness in the task manipulation used in studies 1 and 2 (see 11.5.1).

More puzzling still is the lack of a divided attention effect in Study 3, where no difference was found between the baseline and divided attention conditions, as predicted by the CLCC model. This seems highly unlikely to be a reflection of a lack of effect, however. Every model reviewed in chapter 3 proposes an important relationship between attention and presence; it is highly unlikely that all of these models are mistaken; furthermore, Study 6 shows that the attention focus DES-II factor is a significant predictor of presence. More likely, a methodological artifact (such as differential task difficulty – see point 2 in 7.9.1, chapter 7, and 11.5.1 below) has lead to the lack of effect in this study.

In general then, Studies 1, 2 and 3 provide moderate support for role of working memory in presence, and inconclusive evidence for the role of attention in presence. More research will be required to close these questions (see 11.5).

11.2 Cognitively higher-level and lower-level forms of presence

Three of the four ITC-SOPI factors show an interesting pattern of results across studies, which suggest that presence is a cognitively complex phenomenon.

Spatial – In Studies 1 and 2, this factor showed no working memory interference effects. In Study 5, the spatial factor showed the lowest R^2 value, indicating that it was the hardest to predict from cognitive factors (see Table 9.8 in chapter 9). Finally, in Study 6, the spatial factor showed no non-diegetic information integration effect. As a whole, this evidence suggests that spatial presence is not significantly affected by semantic (content) information, as argued by Slater (2003a), and suggests that it is largely perceptual, relying on highly optimized, low-level cognitive processes which require minimal resources (although see 11.2 above)

Engagement – In Study 1, engagement showed a small interference effect when the verbal working memory system was loaded, suggesting that engagement requires some semantic integration (the effect was not replicated in Study 2). In Study 5 the role of semantic information in engagement became clearer – it shows strong expertise effects in the predicted direction (more expertise leading to lower engagement). However, engagement showed no effect for integration of non-diegetic information in Study 6, although it was predicted by change in emotion. These results suggest that engagement, as proposed by the CLCC model, does involve higher-level semantic processing. This processing seems to involve the integration of diegetic VE information with existing knowledge and expectations of the VE content.

Naturalness – The evidence suggests that this factor is more closely related to engagement than to spatial presence. In Study 1 naturalness also shows small interference effects when spatial working memory was loaded (this was not replicated in Study 2). Study 5 showed clear high-level processing effects, although the pattern was slightly different from that showed by engagement – content expertise does not reduce naturalness, and priming is a factor. This suggests that naturalness also operates on the basis of expectation, but rather than being specifically about semantic information, is about expectation in the most general sense (in terms of interaction, content, etc). Study 6 confirmed this by showing that non-diegetic information (which creates extremely diffuse expectations of the content) affected naturalness in the predicted direction (see 11.5 below).

Negative effects - The negative effect factor is slightly problematic in this context, as it is not directly considered part of presence, but rather a negative correlate of it

(Lessiter *et al.*, 2001). Nonetheless, following the expectation matching feature of the CLCC model (which was demonstrated in Study 5), considering negative effects (particularly simulator sickness), may lead to interesting insights. Kennedy *et al.* (1988) have argued that simulator sickness arises due to a lack of match between sensory stimulation on various sense organs; in effect, stimuli at one organ create an expectation, which if not matched by the other organs, leads to simulator sickness. Comparing the negative effects findings then might give some insight on this expectation hypothesis. In Studies 1, 2 and 3, no working memory or attention effects were found, as one would expect, due to the low level, sensory nature of these negative effects. In studies 5 and 6, which deal more explicitly with expectation, some weak effects are found on this factor (priming and thematic inertia were predictors), suggesting that some complex interactions may occur. Nonetheless, these findings should be interpreted with care, as the manipulations used were not at the correct level of cognitive abstraction for testing these negative effects appropriately.

One can therefore conclude from these findings that that presence occurs at various levels of abstractness, which in turn strongly suggests that presence is the result of complex information processing (Craik & Lockhart, 1972). Spatial presence represents the most concrete of these levels, being hardly affected by semantic information and expectation. Next is naturalness, which shows small semantic effects, relying more on general expectations based both on semantic knowledge and previous experiences with similar media; and most abstract of the three is engagement, which relies heavily on semantic information and is subject to interference from semantic expertise. Furthermore, it seems that spatial presence makes use of dedicated cognitive processes which make use of very little working memory (an amount not detectable by the methods used in these studies), while engagement and naturalness require some working memory to maintain and test expectations. This adds support to the idea that spatial presence is associated with specialized neural circuitry (Sanchez-Vives & Slater, 2005).

This hierarchy of abstractness is reminiscent of the LOP model of presence (Riva & Waterworth, 2003), with two important differences: One, the CLCC model does not differentiate between proprioceptive and perceptual levels (as it is an computational rather than an existential model); and two, the CLCC model is unique in that it supports a *continuum* of contributions of perceptual and conceptual information rather than a three level hierarchy. The existence of three levels in this discussion is actually an artifact brought about by the three factor scale used in the studies. It is therefore incorrect to refer to different set types of presence; rather, the CLCC model supports the existence of cognitively higher and lower forms of presence (i.e. forms which make use of more or less conceptual information).

11.3 'Presence' or spatial presence?

Slater (2003a) explicitly identified all forms of non-spatial presence as not being presence, but rather correlates of presence (see 2.4 in chapter 2 for this argument). The question is then, should engagement and naturalness be considered to be presence, or should they be excluded from the core of presence theory and considered only as correlates? In an important sense, Slater's argument cannot be opposed, because it is not a scientific argument; it is semantic. There is little theoretical reason to consider spatial presence as special – of all the models reviewed in chapter 3, only the MEC model considers spatial presence separately to other forms, and this was done for pragmatic rather than scientific reasons (Wirth *et al.*, 2007). Similarly, there is no published empirical evidence which shows that manipulations of presence lead to no changes in engagement, or vice versa. It is therefore not currently possible to separate spatial presence as being the basic process which produces the other forms of presence. It is equally likely that all forms of presence (including spatial) are the products of other, far more fundamental psychological processes (as proposed by the CLCC model).

In order to separate spatial presence from other forms of presence, one would have to show that psychologically they are distinct but related processes (perhaps by finding a clinical double dissociation, or by means of brain imaging), in the same way that *thinking about walking* and *walking* are distinct but related processes. The CLCC model was built under the assumption that spatial presence is not unique, but rather is part of a set of emergent properties which lie on a continuum of abstractness, which arise due to information processing. The existence of the information processing mechanisms involved is supported by the evidence discussed in 11.2 above.

11.4 The role of expectation in presence

The CLCC model proposes that subjects' expectations (in the form of the pervasive semantic bias) play a central role in presence. The data from Studies 5 and 6 strongly support the content expectation principle stated in section 9.9.1 (chapter 9): VE relevant knowledge creates a cognitive context of expectations, with more knowledge leading to more specific expectations; and presence is more likely to occur when expectations are matched by the VE system. Furthermore, the fact that only the higher cognitive forms of presence use working memory (Study 1), suggests that this comparison requires a certain degree of processing.

There is one important caveat to this principle. The evidence does not show that expectation is a necessary condition for presence; strictly speaking, expectations are only correlates of presence. Although expectation was manipulated in Study 6, the size of the expectation effect was modest, and the baseline condition showed significant levels of presence, even in the absence of a particular expectation. One could attempt to create a VE which creates no expectations at all, to see if presence is possible without expectation. However, associative models of human memory (Rumelhart *et al.*, 1986; Rumelhart & Ortony, 1977) suggest that an expectation will always occur as even small stimuli can become powerful cues for retrieval of previous similar contexts. It may therefore be impossible to test a complete absence of expectation, although it should still be possible to run studies which create one particular expectation, and then violate it by degrees; the CLCC prediction would be that the wider the discrepancy between the expectation and the stimuli, the more it would reduce presence; but it is not possible to determine what would occur in the absence of expectation (if such a condition is even possible).

A possible exception to the expectation principle may be spatial presence. Expectation effects are most clear for naturalness (Studies 1, 5 and 6), and engagement (Studies 1 and 5). However, spatial presence only shows expectation effects Study 5. Given that Study 5 used a relational design, one must conclude that currently there is little evidence for the role of expectation in spatial presence (see 11.5.2).

11.5 Further development of the CLCC model

11.5.1 Further investigations into working memory and attention

As discussed in 11.1, some methodological artifact may have masked working memory effects in Studies 1 and 2. Future studies should attempt to overcome these problems and conclusively establish working memory effects. Two possible modifications to the current method are suggested: One, working memory load should be increased to ensure unambiguous quantification of the predictor, which would increase the power of the between-groups comparison; this is particularly important for the verbal loading task, which seemed to be too weak (see 6.8.2 in chapter 6). Two, the task used only reliably loaded subjects during some parts of their experience (when moving between code pad and door). It is possible that subjects had enough unloaded time in the VE to affect their scores. A new task should be developed which ensures that subjects are loaded during the entire experience.

Similarly, the attention manipulation used in Study 3 could have introduced a significant source of error. The attention loading task (keypress on seeing the blue rectangle appear) may have required too little attention to complete, preventing it from successfully loading attention. A less attractive stimulus should be chosen (a smaller rectangle, or a colour with less contrast relative to the background) to increase the amount of attention required to respond. The degree of attractiveness should also be established empirically before use.

11.5.2 Further investigations into expectation effects

The data seem to suggest a significant role for expectation in the cognitively higher forms of presence (see 11.4 above). The most solid evidence comes from Study 5, which used only a single content area, unfortunately limiting the generalizability of its findings. It would therefore be useful to run a repetition of Study 5, using several randomly selected content areas in a longitudinal study. Subjects who know nothing of the selected content areas would be taken to set levels of expertise, and then have run through content related VEs to determine the effect of expertise. It would also be useful to repeat Study 6, using a set of randomly assigned content areas, together with baseline measures of subject expertise to control for pre-existing knowledge.

11.5.3 Emotion and presence

Although emotion was included only as a control in Study 6, the findings raise interesting questions about the relationship between presence and emotion. The findings generally agree with the predictions of Alcañiz *et al.* (2003) and Baños *et al.* (2004), which suggests that there is some real relationship between presence and emotion. The CLCC model does not make any predictions about emotion, as it is an information processing model only; it assumes that all meaning (be it semantic, emotional or otherwise) is equivalent. It may be the case, as suggested by Damasio (1999), that emotion is special in that it allows for assigning rewards and punishments to ensure expectations match actual outcomes (O'Doherty *et al.*, 2001), or by adjusting the locus of attention (Rich *et al.*, 2005). It would be beneficial to consider the inclusion of emotion as a system-wide moderating force, once more empirical evidence for the relationship between presence and emotion becomes available.

Appendix A: The Independent Television Commision's Sense of Presence Inventory (ITC-SOPI)

A.1 Overview

This instrument is a cross-media measure which has been thoroughly evaluated in terms of validity and reliability (Lessiter *et al.*, 2001). The scale was developed by factor-analyzing 63 Likert-type items created from a review of the literature, which led to four factors (in decreasing eigenvalue order):

1. *Sense of physical space (spatial presence):* The degree to which the subject has a sense of being in the space of the VE, and that the objects and characters in the VE occupy the space as the subject.

2. *Engagement:* A sense of psychological involvement with and enjoyment of the VE content.

3. *Naturalness (Ecological validity):* The sense that the VE and its content are lifelike or realistic.

4. *Negative effects:* Measures negative physiological effects (such as dizziness and eyestrain) – this factor is negatively correlated with the other three factors.

The final form of the scale retained only 44 of the original 63 items over the four factors (physical space: 19 items; engagement: 13 items; naturalness: 5 items; negative effects: 6 items). The four factors are conceptually independent, so that a single presence value cannot be produced by the scale – rather, each measure produces four independent values which are supposed to measure separate aspects of the experience (although in practice the first three factors often correlate significantly with each other (Lessiter *et al.*, 2001; Nunez & Blake, 2006). Further details of this scale can be found in section 2.4.1.4 in Chapter 2.

A.2 Presentation of items

In all studies reported, the items were presented in the order given by Lessiter *et al.* (2001), namely: a1, a2, a3, a4, a5, a6, b1, b2, b3, b4, b5, b7, b8, b9, b10, b11, b12, b13, b14, b15, b16, b17, b18, b19, b20, b21, b22, b24, b25, b26, b27, b28, b29, b30, b31, b32, b33, b34, b35, b36, b37, b38.

All items were presented with a seven point Likert response scale, anchored by "Strongly disagree" on the left (scoring 1) and "Strongly agree" on the right (scoring 7), as suggested by Lessiter *et al.* (2001).

A.3 Items in each factor

A.3.1 Sense of physical space (spatial presence)

Item number	*Item stem*
b12	I felt I wasn't *just* watching something.
b13	I had the sensation that I moved in response to parts of the displayed environment
b18	I had a sense of being in the scenes displayed.
b19	I felt that I could move objects (in the displayed environment).
b22	I could almost smell different features of the displayed environment.
b24	I had a strong sense of sounds coming from different directions within the displayed environment.
b25	I felt surrounded by the displayed environment
b28	I felt I could have reached out and touched things (in the displayed environment)
b29	I sensed that the temperature changed to match the scenes in the displayed environment.
b31	I felt that *all* my senses were stimulated at the same time.
b33	I felt able to change the course of events in the displayed environment.
b34	I felt as though I was in the same space as the characters and/or objects.

Item number	***Item stem (continued from previous page)***
b35	I had the sensation that parts of the displayed environment (e.g. characters or objects)were responding to me.
b36	It felt realistic to move things in the displayed environment.
b38	I felt as though I was participating in the displayed environment.
b4	I felt I could interact with the displayed environment.
b7	I felt that the characters and/or objects could almost touch me.

A.3.2 Engagement

Item number	***Item stem***
a1	I felt sad that my experience was over
a3	I had a sense that I had returned from a journey
a4	I would have liked the experience to continue
a5	I vividly remember some parts of the experience
a6	I'd recommend the experience to my friends.
b1	I felt myself being 'drawn in'.
b16	My experience was intense.
b17	I paid more attention to the displayed environment than I did to my own thoughts (e.g., personal preoccupations, daydreams etc.).
b2	I felt involved (in the displayed environment).
b3	I lost track of time.
b30	I responded emotionally
b32	The content appealed to me.
b8	I enjoyed myself.

A.3.3 Naturalness (Ecological Validity)

Item number	***Item stem***
b11	The content seemed believable to me.
b15	I felt that the displayed environment was part of the real world.
b20	The scenes depicted could really occur in the real world
b27	I had a strong sense that the characters and objects were solid.
b5	The displayed environment seemed natural.

A.3.4 Negative effects

Item number	***Item stem***
a2	I felt disorientated
b10	I felt tired.
b14	I felt dizzy.
b21	I felt I had eyestrain.
b26	I felt nauseous.
b37	I felt I had a headache.

Appendix B: The Differential Emotion Scale, Second Edition (DES-II)

B.4

This is the second edition of Izard's differential emotions scale, the DES-II (Izard, 1991). This scale contains 30 items, each either a word or phrase describing an emotion; subjects are asked to rate the degree to which they felt that emotion during the experience on a seven point scale (1 = Not at all, 7 = Very much). The Izard DES-II has been recently validated and psychometrically evaluated for research use (Fuenzalida *et al.*, 2005). The DES-II was used in Study 6 (see chapter 10).

B.5 Presentation

The DES-II was presented electronically. Subjects were shown the instruction "*During your experience in the displayed environment, did you feel...*" underneath which appeared the item, and under that seven checkboxes for response. Subjects chose when to see the next item (by clicking a "next" button), but could not go back to previously completed items. The order of item presentation was randomized for each subject.

B.6 Items

The following 30 items comprise the DES-II. The thirteen highlighted items represent those used in Study 6, following the factor analysis (those with a factor loading higher than 0.7).

The positive emotion factor consisted of: *joyful, happy, surprise, amazed*

The negative emotion factor consisted of: *disgusted, disdainful, downhearted, angry, a feeling of revulsion, scornful.*

The attention focus factor consisted of: *alert, you were concentrating, attentive*

The following items did not load on any factor: *a feeling of distaste, blameworthy, enraged, guilty, fearful, sheepish, delighted, astonished, discouraged, shy, afraid, mad, scared, bashful, contemptuous, sad, repentant.*

Appendix C: Gamer's survey used in Study 4

C.1 Presentation of items

The order of all the items was randomized for each subject. The entire survey was presented on one web page, and subjects could not submit their data until all items were completed. All items were presented in a seven point Likert format, with 'not at all' and 'very much so' as the anchors.

C.2 Items in each factor

C.2.1 Length of time playing presence games

How long have you been playing the following types of game?

Item	*Response categories*
Simulators (e.g. Pacific Fighters, Need for Speed, Microsoft Flight Simulator, etc.).	Less than 1 year / 1-2 years / 2-5 years / More than 5 years
First person shooter (e.g. Half-Life, Counterstrike, Far Cry, DOOM 3, etc.).	
RPGs (e.g. Neverwinter Nights, Bard's Tale, Everquest, etc.).	

C.2.2 Frequency of playing presence games

How often do you play the following types of game?

Item	*Response categories*
Simulators (e.g. Pacific Fighters, Need for Speed, Microsoft Flight Simulator, etc.).	Never / A few times a month / a few times a week / almost every day
First person shooter (e.g. Half-Life, Counterstrike, Far Cry, DOOM 3, etc.).	
RPGs (e.g. Neverwinter Nights, Bard's Tale, Everquest, etc.).	

C.2.3 Length of time playing non-presence games

How long have you been playing the following types of game?

Item	*Response categories*
Fighting games (e.g. Tekken, Soul Calibur, Street Fighter, etc.).	Less than 1 year / 1-2 years / 2-5 years / More than 5 years
Real-time strategy (e.g. Command & Conquer Generals, Dawn of War, etc.).	
Abstract puzzles (e.g. Tetris, Solitaire, Puyo-Puyo Fever, Bust-a-move, etc.).	

C.2.4 Frequency of playing presence games

How long have you been playing the following types of game?

Item	*Response categories*
Fighting games (e.g. Tekken, Soul Calibur, Street Fighter, etc.).	Less than 1 year / 1-2 years / 2-5 years / More than 5 years
Real-time strategy (e.g. Command & Conquer Generals, Dawn of War, etc.).	
Abstract puzzles (e.g. Tetris, Solitaire, Puyo-Puyo Fever, Bust-a-move, etc.).	

C.2.5 Knowledge of media

These two factors had a single item each:

Item	*Response categories*
How much knowledge do you have about how computers work? (select the word that best describes you)	Basic / Intermediate / Expert
How much knowledge do you have about how computer games work? (select the word that best describes you)	

C.2.6 Integration of non-diegetic information

Item
The background music of a game enhances the game experience for me.
I find that bad dialogue or a poor storyline ruins the game experience for me.
For me, the story/plot of the game is an important part of the experience.
Inappropriate music in a game can ruin the game experience for me.
It is important for me that the story/plot in a game be consistent with the game's world.

C.2.7 Self-rated importance of presence

Item
I prefer a game which is realistic over one which is abstract.
A game should make me feel as if I am transported to inside the game world.
I prefer games which create a sense of being in a place.
For me, the most important aspect of game playing is the ability to explore other worlds.
The quality of a game's graphics are very important for my game experience.
The quality of a game's sound are very important for my game experience.

C.2.8 Thematic inertia

Item
I prefer playing games which are related to my other hobbies and interests.
After playing a game, I often want to play more games in the same setting/genre.
Reading the manual or visiting the website of a game puts me in the mood for playing the game.
After watching a TV program or film, I often feel like playing a game that is similar to the film or program.
After playing a game, I feel I want to know more about the setting/genre of the game.
I choose games based mostly on their setting/genre

C.2.9 Presence management strategies

Item
When I play, I turn off the lights and try to keep the room dark.
As far as I can afford it, I make sure my computer has the best hardware for playing games.
If I am disturbed while I am playing, it ruins the experience for me.
When I play, I try to minimize distractions.
I try to get the latest games to play.
I will consider upgrading my computer to play a particular game.

Appendix D: Flight simulator users' survey used in Study 5

D.1 Presentation of items

The order of all the items was randomized for each subject (the exception to this were the age and gender items, which were always present first, in that order). The entire survey was divided into five pages of equal size (number of items) to speed up loading times, and to prevent subjects from loosing their position in the survey. Each of the five pages of the survey required the completion of all items before continuing to the next page. The items of this survey were then followed by the ITC-SOPI (see Appendix A). All items were presented with a semantic differential response scales. The anchors for each item are presented in the tables below. In some cases, each point in the scale was labeled for clarity.

D.2 Items in each factor

D.2.1 Thematic Inertia (Cronbach's alpha = 0.79)

Item	***Response scale / anchors***
Reading about real world aviation or flight in a book, magazine or web-page makes me want to play a flight simulator.	Very rarely / almost always
During a flight in a real aircraft, I feel like playing a flight simulator.	Very rarely / almost always
After watching a film, documentary or television programme about aviation or flight, I feel like playing a flight simulator.	Very rarely / almost always
Item	***Response scale / anchors (continued from previous page)***

After browsing through flight simulator add-ons (aircraft, sceneries, missions, flights, etc.), I feel like playing a flight simulator.	Very rarely / almost always
After going to an airport, I feel like playing a flight simulator.	Very rarely / almost always

D.2.2 Content knowledge (Cronbach's alpha = 0.63)

Item	***Response scale / anchors***
I prefer to fly virtual flights around places which I have been to in real life.	Very rarely / Almost always
I tend to fly around the same small set of places each time I play a flight simulator.	Very rarely / Almost always
I tend to use the same small set of airplanes each time I play a flight simulator.	Very rarely / Almost always
I like to learn about the airplanes which I fly in my flight simulators.	Very rarely / Almost always
I like to learn about the places which I fly around in my flight simulators.	Very rarely / Almost always
I prefer to fly airplanes in flight simulators which I already know a lot about.	Very rarely / Almost always
I tend to fly missions or flights similar to real world flights I have been on	Very rarely / Almost always

D.2.3 Enjoyment (Cronbach's alpha = 0.68)

Item	***Response scale / anchors***
In general, I enjoy playing commercial flight simulators (Microsoft Flight Simulator, FLY!, Flightgear, etc.)	I do not enjoy them at all / I enjoy them very much
In general, I enjoy playing combat flight simulators (Microsoft Combat Flight Simulator, Lock On, Pacific Fighters, Falcon 4, etc.)	I do not enjoy them at all / I enjoy them very much
I normally find playing commercial flight simulators (Microsoft Flight Simulator, FLY!, Flightgear, etc.) to be a fun experience.	I do not find it a fun experience / I find it a fun experience
I normally find playing combat flight simulators (Microsoft Combat Flight Simulator, Lock On, Pacific Fighters, Falcon 4, etc.) to be a fun experience.	I do not find it a fun experience / I find it a fun experience
I look forward to playing commercial flight simulators (Microsoft Flight Simulator, FLY!, Flightgear, etc.).	I do not look forward to it / I look forward to it very much
I look forward to playing combat flight simulators (Microsoft Combat Flight Simulator, Lock On, Pacific Fighters, Falcon 4, etc.)	I do not look forward to it / I look forward to it very much

D.2.4 Evaluation of simulator realism (Cronbach's alpha = 0.76)

Item	***Response scale / anchors***
In general, I find current commercial flight simulators (Microsoft Flight Simulator, FLY!, Flightgear, etc.) offer a realistic experience.	I do not find the experience realistic at all / I find the experience quite realistic.
In general, I find current combat flight simulators (Microsoft Combat Flight Simulator, Lock On, Pacific Fighters, Falcon 4, etc.) offer a realistic experience.	I do not find the experience realistic at all / I find the experience quite realistic.
I consider combat flight simulators (Microsoft Combat Flight Simulator, Lock On, Pacific Fighters, Falcon 4, etc.) to be computer games rather than simulations.	I consider them to be computer games / I consider them to be simulations.
I consider commercial flight simulators (Microsoft Flight Simulator, FLY!, Flightgear, etc.) to be computer games rather than simulations.	I consider them to be computer games / I consider them to be simulations.
The experience provided by current commercial flight simulators (Microsoft Flight Simulator, FLY!, Flightgear, etc.) is like the real thing.	It is not like the real thing at all / It is very close to the real thing
The experience provided by current combat flight simulators (Microsoft Combat Flight Simulator, Lock On, Pacific Fighters, Falcon 4, etc.) is like the real thing.	It is not like the real thing at all / It is very close to the real thing

Item	***Response scale / anchors (continued from previous page)***
When I play a flight simulator, I use real flying procedures and restrictions as much as the game permits.	Very rarely / almost always
When I play a flight simulator, I make use of the "time compress" feature.	Very rarely / almost always
When I play a flight simulator, I adjust the game's weather and time of day to match the real-world weather and time as much as the flight simulator allows.	Very rarely / almost always
When I play a flight simulator, I will sometimes disable or ignore features of the game so that I can more closely follow real world flying procedures.	Very rarely / almost always

Appendix E: VE scenario instructions

E.1 Presentation of instructions

In each case, the instructions were printed on an A4 sized paged, in Arial 18 point font. No heading of other text was presented with the scenario instructions. Subjects were given sufficient time to read the instructions twice.

The hospital scenario was used in Study 1 and Study 2 (see chapters 5 and 6 respectively). The monastery scenario was used in Study 6 (see chapter 10).

E.2 Hospital scenario

In this scenario, you are a builder at a hospital which is under construction. While working late one night, you discovered that you have been left behind – everyone has gone and the building is locked up!
You only have a key to get out through the parking garage, which is in the basement of the hospital – you are now on the top floor. To get out you will need to make your way to the bottom floor, and find the parking garage.
The hospital's security system is already working, so you will come across closed security doors. The codes to open the doors can usually be sound somewhere near to the door.

E.3 Monastery scenario

In this scenario, you are a research assistant to a history professor. She has negotiated with the monks of a nearby monastery to take their collection of ancient books for study.
She has asked you to go to the monastery and collect the books. The monks are away at a retreat, but they have told you that you should feel free to wander the monastery collecting the books, and enjoying the architecture and the grounds.
The monastery has two floors, a basement, and there is a chapel behind the main building. You need to search all these places for the books.

Appendix F: Priming materials used in Study 6

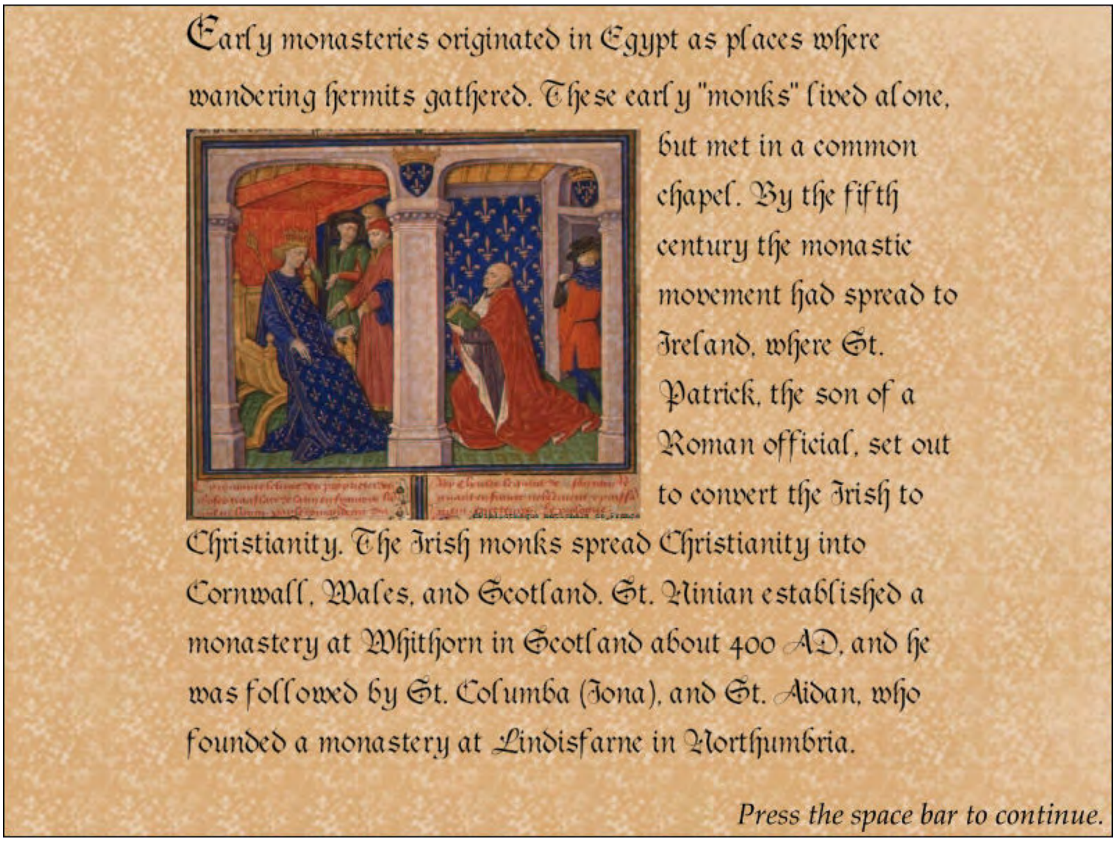
F.1 Overview

These materials are based on those used in Nunez & Blake (2003a). The two sets (monastery-relevant and monastery-irrelevant) were matched in terms of word count.

F.2 Presentation

The priming materials were arranged as a slide show, and presented electronically on the same workstation as was used for the VE experience. The subjects could click to move forward to the next slide, but could not go back to previous slides. They were instructed to slowly read the text and look at the images, while allowing themselves to think about the content. The slides were presented for three timed minutes, and then removed.

F.3 Monastery-relevant slides (VE-relevant)



These Celtic monasteries were often built on isolated islands, as the lifestyle of the Celtic monks was one of solitary contemplation.

The big change in this early monastic existence came with the establishment of the "Benedictine Rule" in about 529 AD. The vision of St. Benedict was of a community of people living and working in prayer and isolation from the outside world. The Benedictine Rule was brought to the British Isles with St. Augustine when he landed in Kent in 597 AD.

Over the next thousand years a wide variety of orders of monks and nuns established communities throughout the British Isles. These orders differed mainly in the details of their religious observation and how strictly they applied those rules. The major orders that established monastic settlements in Britain were the Benedictines, Cistercians, Cluniacs, Augustinians, Premonstratians, and the Carthusians.

Press the space bar to continue.

The first buildings of a monastic settlement were built of wood, then gradually rebuilt in stone. The first priority for rebuilding in stone was the chancel of the church. This way of proceeding meant that the rest of the monastery was at risk of fire, which accounts for the fact that many of the monastic remains you can visit today are in the later Gothic style of architecture.

Typically a monastery was in the charge of an abbot. The abbot was responsible for the souls of the monks, which often

Press the space bar to continue.

meant he was responsible for imposing the rules of the order upon the monks. This included having the power to beat or to imprison in chains. The abbot's deputy was the prior, the person most likely to carry out the disciplinary actions. Monasteries also had an almoner, responsible for the distribution of charity (food and clothing) to the poor. A cellarer was responsible for supplies of food and drink, a sacrist looked after the church and lay servants were employed by monks as the monastic houses became wealthier.

Press the space bar to continue.

Abbeys grew their own food, did all their own building, and in some cases, grew quite prosperous doing so. Fountains Abbey and Rievalux, both in Yorkshire, grew to be enormously wealthy, largely on the basis of raising sheep and selling the wool.

Throughout the Dark Ages and Medieval period the monasteries were practically the only repository of scholarship and learning. The monks were by far the best educated members of society - often they were the only educated members of society. Monasteries acted as libraries for ancient manuscripts, and many monks were occupied with laboriously copying sacred texts (generally in a room called the scriptorium).

Press the space bar to continue.

On another front this was also the period of manuscript books, works produced by the monks showing illuminated calligraphy. In the areas where Celtic influence was strongest, for example in Northumbria, the monks created "illuminated" manuscripts; beautifully illustrated Bibles and prayer books with painstakingly created images on most pages. These illuminated manuscripts, such as the Lindisfarne Gospel, are among the most precious remnants of early Christian Britain.

Press the space bar to continue.

One of the main sources of revenue for monasteries throughout the medieval period were pilgrims. Pilgrims could be induced to come to a monastic house by a number of means, the most common being a religious relic owned by the abbey. Such a relic might be a saint's bone, the blood of Christ, a fragment of the cross, or other similar religious artefact. The tomb of a particularly saintly person could also become a target for pilgrimages. Pilgrims could generally be induced to buy an isignia which proved they had visited a particular shrine. Some popular pilgrimage centres built hotels to lodge pilgrims.

Press the space bar to continue.

F.4 Monastery-irrelevant slides (VE-irrelevant)

The most spectacular part was climbing into the firebox of a Bulleid Pacific, two people at a time. The trick is to grasp a fairly high-up handle in each hand, swing both feet into the fire-door opening, transfer hand grip to lower handles, ease body further in, turn over and wriggle the remaining distance. Inside the firebox we could easily identify the components, including the enormous thermic siphons. Coming out of the firebox was a slight variation on the entry procedure: squeeze out, roll over and get two people to help you up. Clive did give us one warning: on these engines, there's a stub lever sticking up beside the firebox: it's used for rocking the grate. Make sure somebody is covering it because if you slip, you'll be like the engine itself: you'll have a tender behind!

Press the space bar to continue.

Another thing we had to learn was the name of each track and which signal controlled which road: which switch was controlled from the box in the station, and which could be accessed by throwing the point lever in the yard, adjacent to the switch. There were the two platform roads in the station, the Pump-house siding, the Newick road (which used to be the running line to Newick when the Bluebell was a "real" railway), the headshunt, and the six yard tracks. There were the starter signals, the two signals controlling entrance to the two station roads and their shorter counterparts allowing cautious entry even when the track was occupied, and "dummy" signals at ground level.

During the second afternoon we were introduced to "our" locomotive, no. 263. It was an 0-4-4 tank engine built in about 1905 for the South-Eastern and Chatham Railway. From our point of view, it had two "interesting" characteristics: it used the regular train vacuum brake rather than steam brakes for the engine itself, and it had a steam-powered reverser. It was parked over the pit, so we could walk down the steps and look underneath at the points that would need lubrication and examine the reverser mechanism and dampers.

At the end of the second day Clive handed us our exam papers, to be handed in by Friday. The cover sheet was a list of safety rules and regulations which we were to sign as "read and understood". Back at Wayside Cottage, I failed to obey the rule "Look out for metal obstructions above your head". I bent over to unlace my safety shoes in the porch, straightened up, hit my head on a metal flower basket, staggered back, and banged into and cracked a window pane.

Press the space bar to continue.

I was number two, so my first task was to oil the inside and underneath stuff: the axle journals, the big end cranks and the oiling points on the trailing 4-wheeled bogie. Clive tossed me a long once-white coat, with the comment "No need to get your overalls dirty!". I put this on, filled up the lubricating can from the large oil-can, and went under. First the big ends: to do this I had to lay a plank across the pit and climb up on it. In this position I was bending right across the axle, but it was reasonably easy to grab each cork with a rag, twist it out, fill up with oil and replace the cork. One bearing took an incredible amount of oil, but none seemed to be leaking out. The other needed hardly any oil.

Then came the wheel bearings, and then off to the other end of the locomotive. The movement of the bogie (truck) is lubricated by four "onions", open-topped onion-shaped steel capsules, which hold the oil that is siphoned onto the actual bearing surfaces by trimmings. The only problem was how to get oil to flow into the onions, since there was no room to hold the oiling can high enough for oil to flow. The solution was to work with a very full can of oil, and then there was just enough gradient for the liquid to flow.

Press the space bar to continue.

John also showed us how to work the feed-water injector. I'm convinced that this is a black art. Turn on the water three-quarters full, turn on the steam, knock the water back a little, then on a little more, and the injector starts. Well, that's how John did it. For me, it was fiddle, fiddle, fiddle with the water control until John gave it one final gentle knock which started the injector every time.

Clive instructed the student driver. "Make a brake" was easy: move the little black handle to turn on steam to the brake ejector to create a vacuum. Then the reverser, tricky, but not a black art because you could see what was happening by looking at the brass pointer. The method is: put the reverser lever to forward or back, and "blip" the steam control. The brass pointer echoes, on a graduated scale, the position of the reversing gear, and if you overshoot you can put the lever the other way and "blip" the steam control again. Finally, put the lever in the middle.

After a few tries at the reverser, the student could check the signal (and get John to check all was clear on his side), toot the whistle, and push the regulator open.

Press the space bar to continue.

Appendix G: Summary table of models reviewed in Chapter 3

Two-pole / Environment selection model *(early form)*	
Notable publications	Slater *et al.,* (1994) Hendrix & Barfield, (1996)
Structure	Subjects exist in a continuum between two poles ('virtual environment' and 'real environment'). The subject's actual state is governed mostly by immersion; high levels of immersion (and a focusing of attention on the display) tend to move the subject towards the 'virtual environment' pole.
Presence in the model	Position on the continuum, as a continuous state.
Empirical evidence	More than 200 independent studies, particularly associated with role of immersion, multimodality and interactivity. Although many of these studies provide positive evidence, a sizable minority (perhaps 25%) fail to replicate the effects.
Resolution of the five problems	Unable to resolve any of the five

Two-pole / Environment selection model *(later form)*	
Notable publications	Slater & Steed (2000); Slater (2002)
Structure	Subjects exist in one of two states ('present in real world' or 'present in virtual environment'). Movement between the poles is uneven; from real to virtual states is not well defined (although it is strongly related to immersion), while movement from virtual to real occurs due to a break in presence (BIP).
Presence in the model	Presence is a binary state, which the subject selects depending on sensory information.
Empirical evidence	All evidence for early form (see above) applies.

	Additionally, BIP (and binary presence state) was demonstrated by Brogni *et al.* (2003) and Vinayagamoorthy *et al.* (2004)
Resolution of the five problems	Capable of explaining the dream-state problem; others not explainable

Three pole model	
Notable publications	Biocca (2003)
Structure	Subjects exist in a space defined by three poles (real environment, virtual environment, mental imagery environment), movement between which is controlled by spatial attention and spatial updating by the subject.
Presence in the model	A continuous state, which can be divided between the three poles. When the state is close to one of the poles, high-undivided presence occurs. Dividion of presence leads to a low-divided presence state.
Empirical evidence	Numerous studies indicating the importance of focused attention in presence; Evidence for mental imagery concept by Baños *et al.* (2005). Support for importance of spatial updating by Barfield and Hendrix (1995) and Slater *et al.* (1995c)
Resolution of the five problems	Completely solves book, physical space and dream state problems. Cannot solve the remaining two.

Focus, Locus, Sensus model	
Notable publications	Waterworth & Waterworth (2001; 2003)
Structure	The subject exists in a space defined by three dimensions: Focus (presence/absence), Locus (virtual environment/real environment), Sensus (conscious/unconscious). The sensus and focus dimensions interact.
Presence in the model	A continuous state; high degrees of presence occur

	when the subject is closest to the *presence*, *conscious* and *virtual environment* pole; otherwise, the subject is in a state of divided presence or absence.
Empirical evidence	Waterworth *et al.* (2002); Waterworth & Waterworth, 2003.
Resolution of the five problems	Completely solves book, physical space and dream state problems. Cannot solve the remaining two.

Levels of presence model	
Notable publications	Riva, Waterworth and Waterworth (2004)
Structure	The subject's self contains three layers (proto-self, core self and extended self), which process proprioceptive, perceptual and conceptual information respectively. Each layer produces its own form of presence. The layers can operate in synchrony, or independently.
Presence in the model	Continuous. Proto-presence is effectively engaging with the world; Core presence is focusing attention on a selected subset of stimuli; Extended presence is a successful comparison of the internal representation of the world (and future projections of it) to the real world. Maximal presence occurs when all layers operate in synchrony.
Empirical evidence	Some immersion related findings can be taken as evidence; Can also explain Waterworth *et al.* (2002); Waterworth & Waterworth, 2003, but no independent empirical validation has been done.
Resolution of the five problems	Completely solves book, physical space

	and dream state problems. Cannot solve the remaining two.

The Measures, Effects, Conditions (MEC) model	
Notable publications	Wirth *et al.* (2007)
Structure	Presence occurs due to a two-stage process. In the first stage, the subject builds spatial situation models (SSMs) of the environment. The success of this construction depends on both subject and media factors. In the second phase, the hypothesis 'The SSM is the primary ego reference frame' is tested perceptually against the available perceptual information. This testing process again depends on both subject and media factors.
Presence in the model	Presence occurs when the SSM encoding the virtual environment is accepted as the primary ego reference frame (PERF). This state is binary by virtue of this acceptance, but is also continuous, as the degree of presence will vary with the quality of the SSM constructed.
Empirical evidence	Construction of SSMs – Gysbers *et al.* (2004). Adoption of SSM as PERF - Böcking *et al.* (2004). Role of subject factors - Sacau *et al.* (2005); Vorderer *et al.* (2004)
Resolution of the five problems	Can explain Book, dream state and virtual stimuli problems. Unable to account for physical reality and inverse presence problems.

Appendix H: Summary of predictions, findings and support for predictions, tabulated by study

Summary of the five studies conducted for this thesis which included direct predictions of presence effects (Study 4 is excluded as it did not make specific predictions, and also did not measure presence directly).

Study 1 (Chapter 5)	
Aim	To determine the contribution of working memory on presence
Sample	177 (age M = 21.3; 78% women; 22% men)
Presence predictions	1. WM load will negatively affect presence 2. Semantic WM load will negatively affect engagement and naturalness more than spatial 3. Visual WM load will negatively affect spatial more than engagement and naturalness
ITC-SOPI findings	*Spatial* – No effect of WM load or system loaded *Engagement* – No effect of WM load; Higher mean score under any spatial load than under semantic load. *Naturalness* – No effect of WM load; Higher mean score under any semantic load than under spatial load *Negative* – No effect of WM load or system loaded
Prediction support	1. Moderate support - Two of four ITC-SOPI factors affected 2. Partial support – System prediction supported, but not load prediction 3. Partial support – System prediction supported, but not load prediction

Study 2 (Chapter 6)	
Aim	To determine if media decoders use more WM than inherent decoders
Sample	89 (age M = 20.6; 57% women; 43% men)
Presence predictions	1. WM load will negatively affect presence 2. WM load will negatively affect engagement and naturalness more than spatial 3. Semantic WM load will affect presence more than loading the spatial system
ITC-SOPI findings	*Spatial* - No effect of WM load or system loaded *Engagement* - No effect of WM load or system loaded *Naturalness* - No effect of WM load or system loaded *Negative* - No effect of WM load or system loaded
Prediction support	1. No support 2. Minimal support – some evidence that direction of effect is as predicted 3. No support

Study 3 (Chapter 7)	
Aim	To determine the relative contribution of WM and attention in presence
Sample	46 (age M = 21.69; 73% women; 27% men)
Presence predictions	1. Concurrent tasks will reduce presence 2. A concurrent WM task will reduce presence more than a divided attention task
ITC-SOPI findings	*Spatial* – No direct effect; Moderation effect on this factor by task performance *Engagement* – No effects *Naturalness* - No direct effect; Moderation effect on this factor by task performance *Negative* – No effects
Prediction support	1. No direct support; some support through task performance

	moderation 2. No support

Study 5 (Chapter 9)	
Aim	To determine the role of thematic inertia and content knowledge on presence
Sample	461 (age M = 31.7; 100% men)
Presence predictions	1. Generalized knowledge will increase presence, while specific knowledge will decrease it 2. Semantic processing contributes to presence at least as much as perceptual processing
ITC-SOPI findings	*Spatial* – Thematic inertia, evaluation of realism, presence management and age were positive predictors; content knowledge was a negative predictor *Engagement* - Thematic inertia, evaluation of realism, presence management and age were positive predictors; content knowledge was a negative predictor *Naturalness* - Thematic inertia, evaluation of realism, presence management, priming and age were positive predictors *Negative* – Weak model fit. Thematic inertia and presence management are positive predictors
Prediction support	1. Data supports prediction directly 2. Data supports prediction, although exact measure of relative contribution is not possible

Study 6 (Chapter 10)	
Aim	To test semantic content manipulation effects on presence
Sample	181 (age M = 21.45; 80% women; 20% men)
Presence predictions	1. Priming will increase presence, more so for engagement and naturalness 2. Semantic fit between non-diegetic music and VE will increase presence, more so for engagement and naturalness 3. Lack of semantic fit between non-diegetic music and VE will lead to no effect on presence
ITC-SOPI findings	*Spatial* – Positive emotion and attention focus are positive predictors, but neither priming nor non-diegetic music condition are predictors. *Engagement* - Positive emotion, negative emotion and attention focus are positive predictors, but neither priming nor non-diegetic music condition are predictors. *Naturalness* - Positive emotion and attention focus are positive predictors. Non-diegetic fit is a positive predictor, with the fit condition increasing mean scores, and the control and non-fit conditions showing no difference. Priming was not a predictor. *Negative* – Negative emotion and priming are predictors (the VE matched priming condition lead to lower mean scores)
Prediction support	1. Weak support – only negative factor is subject to priming (not engagement or naturalness as predicted) 2. Moderate – the naturalness factor showed the predicted effects, but engagement did not 3. Moderate – when non-diegetic music was a predictor, the non-fit condition had no effect as predicted.

References

Alcañiz, M., Baños, R., Botella, C., & Rey, B. (2003). The EMMA project: Emotions as a determinant of presence. *Psychnology, 1*(2), 141-150.

Allen, R. C., & Singer, M. J. (2001). *Presence in altered environments: Changing parameters and changing presence*. Orlando, FL.: U.S. Army Research Institute for the Behavioural and Social Sciences.

Allmendinger, L. (1998). *Diagrammatic literacy.* Paper presented at the Vision Plus 5. Learning meets information design, information design meets Learning.

Anastasi, A., & Urbina, S. (1996). *Psychological Testing*. New York, NY.: Prentice Hall.

Andreasen, N. C., O'Leary, D. S., Arndt, S., Cizadlo, T., Hurtig, R., Rezai, K., Watkins, G. L., Ponto, L. B., & Hichwa, R. D. (1996). Neural substrates of facial recognition. *The Journal of Neuropsychiatry and Clinical Neurosciences, 8*(2), 139-146.

Atkinson, R., & Shiffrin, R. (1968). Human memory: A proposed system and its control processes. In K. Spence & J. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York, NY: Academic Press.

avsim.com. (2003). The AVSIM 2003 survey.

Axelsson, A., Abelin, A., Heldal, I., Schroeder, R., & Widestrom, J. (2001). Cubes in the cube: A comparison of a puzzle-solving task in a virtual and a real environment. *Cyberpsychology & Behaviour, 4*(2), 279-286.

Baddeley, A. D. (1986). *Working Memory*: Claredon Press.

Baddeley, A. D. (1998). Recent developments in working memory. *Current Opinions in Neurobiology, 8*(2), 234-238.

Baddeley, A. D. (2004). *Your Memory: A User's Guide*. Richmond Hill, Ontario: Firefly books.

Baddeley, A. D., Chincotta, D., Stafford, L., & Turk, D. (2002). Is the word length effect in STM entirely attributable to output delay? Evidence from serial recognition. *Quarterly journal of experimental psychology A., 55*(2), 353-369.

Baddeley, A. D., & Liebermann, K. (1980). Spatial working memory. In R. S. Nickerson (Ed.), *Attention and performance* (pp. 521–539). Hillsdale, NJ: Erlbaum.

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2001). Equilibrium revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators and Virtual Environments, 10*, 583-598.

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin, 29*, 1-15.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Baños, R., Botella, C., Cuevas, B., Liaño, V., Alcañiz, M., & Rey, B. (2004). *Mood Induction Procedures Via Virtual Reality: Induction and Change of Sadness and Joy.* Paper presented at the 38th Annual Association for Advancement of Behavior Therapy Convention, New Orleans, USA.

Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., & Rey, B. (2004). Immersion and Emotion: Their Impact on the Sense of Presence. *Cyberpsychology & Behaviour, 7*(6), 734 -741.

Baños, R. M., Botella, C., Guerrero, B., Liano, V., Alcañiz, M., & Rey, B. (2005). The third pole of the sense of presence: Comparing virtual and imagery spaces. *PscyhNology, 3*(1), 90-100.

Barfield, W., Baird, K. M., & Bjorneseth, O. J. (1998). Presence in virtual environments as a function of type of input device and display rate update. *Cyberpsychology & Behaviour, 2*(4), 283-292.

Barfield, W., & 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., & Weghorst, S. (1993). The sense of presence within virtual environments: A conceptual framework. In G. Salvendy & M. Smith (Eds.), *Human-Computer Interaction: Software and Hardware Interfaces* (pp. 699-704): Elsevier.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* Cambridge, MA.: MIT Press.

Bendiksby, M. S., & Platt, M. L. (2006). Segregating the effects of motivation and reflexive visual attention on neuronal activity in area LIP. *Journal of Vision, 6*(6), 744.

Bever, T. (1988). A cognitive theory of emotion and aesthetics in music. *Psychomusicology, 7*(2), 165-175.

Biocca, F. (1997). Cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer Mediated-Communication, 3*(2).

Biocca, F. (2002). *Presence working group research targets.* Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Biocca, F. (2003). *Can we resolve the book, the physical reality and the dream state problems? From the two-pole to a three-pole model of shifts in presence.* Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Bischof, W. F., & Boulanger, P. (2003). Spatial navigation in virtual reality environments: An EEG analysis. *Cyberpsychology & Behaviour, 6*(5), 487-495.

Böcking, S., Gysbers, A., Wirth, W., Klimmt, C., Hartmann, T., Schramm, H., Laarni, J., Sacau, A., & Vorderer, P. (2004). *Theoretical and empirical support for distinction between components and conditions of spatial presence.* Paper presented at the 7th international workshop on presence (PRESENCE2004), Valencia, Spain.

Bordwell, D. (1989). A case for cognitivism. *Iris, 9*, 11-40.

Botella, C., Baños, R. M., & Alcañiz, M. (2003, 6-8 October 2003). *A psychological approach to presence.* Paper presented at the 6th International Workshop on Presence, Aalborg, Denmark.

Bracken, C. C. & Botta, R. A. (2002). *Presence and television: Form vs. Content.* Paper presented at the The fifth international workshop on presence (PRESENCE2002), Porto, Portugal.

Bracken, C. C. (2005a). *Are we together? Exploring the similarities between the concepts of transportation imagery model and presence and their possible impact on persuasion.* Paper presented at the The eigth international workshop on presence (PRESENCE2005), London, UK.

Bracken, C. C. (2005b). Presence and image quality: The case for high-definition television. *Media Psychology, 7*, 191-205.

Bracken, C. C., & Skalski, P. (2006). *Presence and video games: The impact of image quality and skill level.* Paper presented at the The 9th international workshop on presence (PRESENCE2006), Cleveland, OH.

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology: General, 47*, 191-196.

Brogni, A., Slater, M., & Steed, A. (2003). *More breaks less presence*. Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Brown, S., Ladeira, I., & Winterbottom, C. (2003). *The Effects of Mediation in a Storytelling Virtual Environment.* Paper presented at the International conference on storytelling 2003

Bruner, J. (1990). *Acts of meaning*. Harvard, CN: Harvard University Press.

Bystrom, K. E., & Barfield, W. (1996). Effects of participant movement affordance on presence and performance in virtual environments. *Virtual Reality: The Journal of the Virtual Reality Society, 2*(2), 206-216.

Bystrom, K. E., & Barfield, W. (1999). Collaborative task performance for learning using a virtual environment. *Presence: Teleoperators and Virtual Environments, 8*(4), 435-448.

Bystrom, K. E., Barfield, W., & Hendrix, C. (1999a). A Conceptual Model of the Sense of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments, 8*(2), 241-244.

Bystrom, K. E., Barfield, W., & Hendrix, C. (1999b). A conceptual model of the sense of presence in virtual environments. *Presence: Teleoperators and Virtual Environments., 8*, 241-244.

Calvert, S. L. (2001). The social impact of virtual reality. In K. Stanney (Ed.), *Handbook of Virtual Environment Technology*. Hillsdale, NJ: Erlbaum.

Carpenter, P. A., & Just, M. A. (1989). The role of working memory in language comprehension. . In D. Klahr & K. Kotovsky (Eds.), *Complex Information Processing: The Impact of Herbert A. Simon.* (pp. 31-68): Lawrence Erlbaum Associates.

Cavanagh, P., & Mather, G. (1989). Motion: the long and short of it. *Spatial vision, 4*, 103-129.

Cho, D., Park, J., Kim, G. J., Hong, S., Han, S., & Lee, S. (2003). *The dichotomy of presence elements: The where and how.* Paper presented at the IEEE Virtual Reality 2003 Conference, Los Angeles, CA.

Cohen, D., & MacKeith, S. (1991). *The Development of Imagination: The Private Worlds of Childhood.* Lodon, UK: Routledge.

Cohn, V., Di Zio, P., & Lackner, J. (1996). Reaching movements during illusory self-rotation show compensation for expected Coriolis forces. *Society for Neuroscience Abstracts, 22*(1), 654.

Coluccia, E., & Louse, G. (2004). Gender differences in spatial orientation: A review. *Journal of Environmental Psychology, 24*(3), 329-340.

Conway, M. A. (2002). *Levels of processing 30 years on: A special issue of memory.* New York, NY.: Taylor & Francis.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-185.

Craik, F., & Lockhart, R. (1972). Levels of Processing : A Framework for Memory Research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Craver-Lemley, C., & Reeves, A. (1992). How visual imagery interferes with vision. . *Psychological Review, 99*, 633-649.
Dalgarno, B., & Scott, J. (2000). Motion control in virtual environments: A comparative study. In V. Paelke & S. Volbracht (Eds.), *Workshop on Usability Centred Design and Evaluation of Virtual 3D Environments*. Padeborn, Germany: University of Padeborn.
Damasio, A. (1999). *The feeling of what happens:Body emotion and the making of consciousness.* San Diego, CA.: Harcourt Brace.
Decoster, J. & Claypool, H. M. (2004). A Meta-analysis of priming effects on impression formation supporting a general model of information biases. *Personality and Social Psychology Review*, *8,* 2-27.
Darken, R. P., Bernatovitch, D., Lawson, J. P., & Perterson, B. (1999). Quantitative measures of presence in virtual environments: The role of attention and spatial comprehension. *Cyberpsychology & Behaviour, 2*(4), 337-347.
Dillon, C., Keogh, E., Freeman, J., & Davidoff, J. (2001). *Presence: Is your heart in it?* Paper presented at the The fourth internal workshop on presence (PRESENCE2001), Temple, PA.
Dinh, H. Q., Walker, N., Hodges, L. F., Song, C., & Kobayashi, A. (1999). *Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments.* Paper presented at the IEEE Virtual Reality 1999, Houston, TX.
Ellis, S. R. (1996). Presence of mind: A reaction to Thomas Sheridan's "Musings on Telepresence". *Presence: Teleoperators and Virtual Environments, 5*(2), 247-259.
Emoto, M., Masaoka, K., Sugawara, M., & Okano, F. (2004, October 13-15, 2004). *Quantitative evaluation of sensation of presence in viewing the "Super Hi-Vision" 4000-scanning-line wide-field video system.* Paper presented at the The seventh internaltional workshop on presence (PRESENCE2004), Valencia, Spain.
Entertainment.Software.Association. (2000). Game Player Data.
Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence on maximal adaptations on task constraints. *Annual Review of Psychology, 47*, 273-305.
Fencott, C. (1999). *Presence and the Content of Virtual Environments.* Paper presented at the 2nd International Workshop on Presence, Essex, U.K.
Floridi, L. (2005). The philosophy of presence: from epistemic failure to successful observation. *Presence: Teleoperators and Virtual Environments, 14*(6), 656 - 667.
Fodor, J. A. (1983). *The modularity of mind*. Boston, MA.: The MIT Press.
Fontaine, G. (1992). Experience of a sense of presence in intercultural and international encounters. *Presence: Teleoperators and Virtual Environments, 1*(4), 482-490.
Fox, N. A., Henderson, H. A., Marshall, P. J., Nichols, K. E., & Ghera, M. M. (2005). Behavioural inhibition: Linking Biology and Behavior within a Developmental Framework. *Annual Review of Psychology, 56*, 235-262
Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General, 119*(1), 63-76.
Freeman, J., Avons, S. E., Meddis, R., Pearson, D. E., & IJsselsteijn, W. A. (2000). Using behavioural realism to estimate presence: A study of the utility of

postural responses to motion stimuli. *Presence: Teleoperators and Virtual Environments, 9*(2), 149-164.

Freeman, J., Lessiter, J., Keogh, E., Bond, F. W., & Chapman, K. (2004). *Relaxation Island: virtual, and really relaxing.* Paper presented at the Seventh International workshop on presence, Valencia, Spain.

Friedman, D., Brogni, A., Antley, A., Guger, C., & Slater, M. (2005). *Sharing and analyzing presence experiments data.* Paper presented at the The eight international workshop on presence (PRESENCE2005), London, UK.

Fuenzalida, C., Emde, R. N., Pannabecker, B. J., & Stenberg, C. (2005). Validation of the Differential Emotions Scale in 613 mothers. *Motivation and Emotion, 5*(1), 37-45.

Gabreilsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In J. A. Sloboda & P. N. Juslin (Eds.), *Music and Emotion: theory and research.* (pp. 223-249). New York, NY: Oxford University Press.

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cognitive Brain Research, 20*, 226-241.

Garden, S., Cornoldi, C., & Logie, R. H. (2002). Visuo-spatial working memory in navigation. *Applied Cognitive Psychology, 16*, 35-50.

Gerrig, R. (1993). *Experiencing narrative worlds*. Boston: Yale University Press.

Girgus, J. J., Rock, I., & Egatz, R. (1977). The effect of knowledge of reversibility on on the reversibility of ambiguous figures. *Perception and Psychophysics, 22*(6), 550-556.

Goldman, A. (1986). *Epistemology and Cognition.* Cambridge, MA.: Harvard University Press.

Graf, P., & Ryan, L. (1990). Transfer-appropriate processing for implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 978-992.

Gysbers, A., Klimmt, C., Hartmann, T., Nosper, A., & Vorderer, P. (2004). *Exploring the book problem: Text design, mental representations of space and spatial presence in readers.* Paper presented at the The 7th international workshop on presence (PRESENCE2004), Valencia, Spain.

Hastie, T., & Tibshirani, R. (1997). Generalized Additive Models. In *Encyclopedia of Statistical Science*. New York, NY: Wiley.

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

Heeter, C. (2003). Reflections on real presence by a virtual person. *Presence: Teleoperators and Virtual Environments., 12*, 335-345.

Held, R., & Durlach, N. I. (1992). Telepresence. *Presence: Teleoperators and Virtual Environments, 1*, 102-112.

Hendrix, C., & Barfield, W. (1996a). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators and Virtual Environments, 5*(3), 274-289.

Hendrix, C., & Barfield, W. (1996b). The sense of presence within auditory virtual environments. . *Presence: Teleoperators and Virtual Environments, 5*(3), 290-301.

Heraz, A., & Frasson, C. (2006). *Hidden Emotions in Iconic Faces.* Paper presented at the Workshop on Robots and Agents. International Conference on Intelligent Tutoring System (ITS), Jhongli , Taiwan.

Hirschfeld, L., & Gelman, S. (1994). *Mapping the Mind: Domain Specificity in Cognition and Culture.* Cambridge: Cambridge University Press.

Hobson, J. A., Pace-Schott, E., & Stickgold, R. (2000). Dreaming and the brain: Towards a cognitive neuroscience of conscious states. *Behaviour & Brain Sciences, 23*(6), 793-1121.

Hoffman, H. G., Hollander, A., Schroder, K., Rousseau, S., & Furness, T. A. (1999). *Physically touching and tasting virtual objects enhances the realism of virtual environments*. from http://www.hitl.washington.edu/publications/r-99-7.

Hoffman, H. G., Hullfish, K., & Houston, S. J. (1995). *Virtual reality monitoring.* Paper presented at the Virtual Reality Annual International Symposium (VRAIS'95).

Hoffman, H. G., Richards, T., Coda, B., Richards, A., & Sharar, S. R. (2003). The Illusion of Presence in Immersive Virtual Reality during an fMRI Brain Scan. *Cyberpsychology & Behaviour, 6*(2), 127-131.

Hopf, J. M., Boehler, C. N., Luck, S. J., Tsotsos, J. K., Heinze, H. J., & Schoenfeld, M. A. (2006). Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings of the National Academy of Sciences of the United States of America, 103*(4), 1053-1058.

Huang, M., & Alessi, N. (1999). Huang, M., & Alessi, N. (1999). Presence as an emotional experience. In J. D. Westwood, H. M. Hoffman, R. A. Robb & D. Stredney (Eds.), *Medicine meets virtual reality: The convergence of physical and informational technologies options for a new era in healthcare.* Amsterdam: IOS Press.

Hullfish, K. (1996). *Virtual reality monitoring: How real is virtual reality?* , University of Washington, Seattle, WA.

Hwang, J., Kim, G. J., & Rizzo, A. (2004). *Space extension: the perceptual presence perspective.* Paper presented at the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry.

igroup.org. (2004). igroup presence questionnaire (IPQ) Download.

IJsselstiejn, W. A., de Ridder, H., Freeman, J., Avons, S. E. & Bowhuis, D. (2001). Effects of steroscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. *Presence: Teleoperators and Virtual Environments, 10,* 298-311.

IJsselsteijn, W. A. (2002). *Elements of a Multi-level Theory of Presence: Phenomenology, mental processing and neural correlates.* Paper presented at the The fifth international workshop on presence (PRESENCE2002), Porto, Portugal.

IJsselsteijn, W. A. (2004). *Presence in Depth*. Eindhoven, The Netherlands: Technische Universiteit Eindhoven.

IJsselsteijn, W. A., De Ridder, H., Freeman, J., & Avons, S. E. (2000). *Presence: Concept, determinants and measurement.* Paper presented at the Proceedings of the SPIE.

IJsselsteijn, W. A., De Ridder, H., Freeman, J., Avons, S. E., & Bouwhuis, D. (2001). Effects of Stereoscopic Presentation, Image Motion, and Screen Size on Subjective and Objective Corroborative Measures of Presence. *Presence: Teleoperators and Virtual Environments, 10*(3), 298-311.

Ikeno, O., & Davies, R. J. (2002). *The Japanese Mind: Understanding Contemporary Culture*. New York: Charles Tuttle.

Insko, B. (2001). *Passive haptics significantly enhances virtual environments.* University of North Carolina at Chapel Hill.

International Society for Presence Research. (2000). The Concept of Presence: Explication Statement.

Izard, C. E. (1991). *The psychology of emotions*. London, UK: Plenum Press.

Johns, C. L., Nunez, D., Daya, M., Sellars, D., Casanueva, J., & Blake, E. H. (2000). The interaction between individuals' immersive tendencies and the sensation of presence in a virtual environment. In J. R. Mulder & R. van Liere (Eds.), *Virtual Environments 2000: Lecture Notes in Computer Science series*. Vienna, Austria: Springer-Verlag.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness.* Cambridge, MA.: Harvard University Press.

Jonides, J. P. (1981). Voluntary versus automatic control over the mind's eye. In J. Long & A. D. Baddeley (Eds.), *Attention and Performance IX.* (pp. 187-204). Hillsdale, NJ: Erlbaum.

Jorgensen, C., Ogden, J., Willis, E. K., Blessing, M., Caudell, K. A., Patrick, G., & Caudell, T. P. (1997). *Locomotion in a virtual environment: Performance measures and physiological responses.* . Paper presented at the Human Factors and Ergonomics Society 41st Annual Meeting.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kalawsky, R. S. (2000). *The validity of presence as a reliable human performance metric in virtual environments.* Paper presented at the The 3rd international workshop on presnece (PRESENCE2000), Delft, The Netherlands.

Kennedy, R. S., Hettinger, L. J., & Lilienthal, M. G. (1988). Simulator Sickness. In G. H. Crampton (Ed.), *Motion and Space Sickness*. Boca Raton, FL: CRC Press.

Kersten, D., Mamassian, P., & Knill, D. C. (1997). Moving cast shadows induce apparent motion in depth. *Perception, 26*(2), 171-192.

Kose, S. (2003). A psychobiological model of temperament and character: TCI. *Yeni Symposium, 41*(2), 86-97.

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin, 129*, 723-746.

Krubitzer, L. (2005). The evolution of visual cortex in mammals: Genetic and epigenetic contributions to the phenotype. *Journal of Vision, 5*(12), 33a.

Laarni, J., Ravaja, N., Saari, T., & Hartmann, T. (2004). *Personality related differences in subjective presence.* Paper presented at the 7th International Workshop On Presence (PRESENCE2004), Valencia, Spain.

LaBerge, S. (1980). Lucid dreaming as a learnable skill: A case study. *Perceptual and Motor Skills, 51*, 1039-1042.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh : The Embodied Mind and Its Challenge to Western Thought*. London, UK: Harper Collins.

Laptaned, U. (2006). *Situation Awareness in Virtual Environments: A Theoretical Model and Investigation with Different Interface Designs.* Paper presented at the IASTED confernece on computers and advanced teachnology in education, Lima, Peru.

Larsson, P., Vastfjall, D., & Kleiner, M. (2001). The actor-observer effect in virtual reality presentations. *Cyberpsychology & Behaviour, 4*(2), 239-246.

Lavigne, F., & Denis, S. (2001). Attentional and Semantic Anticipations. *International Journal of Computing Anticipatory Systems, 8*, 74-95.

Lee, K. M. (2004). Why presence occurs: Evolutionary psychology, media equation and presence. *Presence: Teleoperators and Virtual Environments., 13*(4), 494-505.

Lee, K. M., & Yung, Y. (2005). Evolutionary nature of the presence experience. *Journal of Cultural and Evolutionary Psychology, 3*, 159-178.

Lee, S., Kim, G. J., & Lee, J. (2004a). *Observing effects of attention on presence with fMRI.* Paper presented at the ACM symposium on Virtual reality software and technology.

Lee, S., Kim, G. J., Rizzo, A., & Park, H. (2004b). *Formation of spatial presence by form or content?* Paper presented at the The seventh international workshop on presence (PRESENCE2004), Valencia, Spain.

Leeb, R., Keinrath, C., Friedman, D., Guger, C., Neuper, C., Garau, M., Antley, A., Steed, A., Slater, M., & Pfurtscheller, G. (2005). *Walking From Thoughts: Not the Muscles Are Crucial, But the Brain Waves!* Paper presented at the The eighth international workshop on presence (PRESENCE2005), London, UK.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments., 10*(3), 282-297.

Lok, B., Naik, S., Whitton, M., & Brooks, F. P. (2003, 22-25 March 2003). *Effects of handling real objects and avatar fidelity on cognitive task performance in virtual environments. .* Paper presented at the IEEE Virtual Reality 2003, Los Angeles, CA.

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

Lombard, M., & Ditton, T. (2004). *A literature-based presence measurement instrument: The Temple Presence Inventory (TPI) (BETA).* Paper presented at the The seventh international workshop on presence (PRESENCE2004), Valencia, Spain.

Lombard, M., Reich, R. D., Grabe, M. E., Bracken, C. C., & Ditton, T. B. (2000a). Presence and television: The role of screen size. *Human Communication Research, 26*, 75-98.

Lombard, M., Ditton, T., Crane, D., Davis, B., Gil-Egui, G., Horvath, K., & Rossman, J. (2000b). *Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil intrucment.* Paper presented at the third international workshop on presence (PRESENCE2000), Delft, The Netherlands.

Lombard, M., & Jones, M. T. (2006). *Defining presence: A framework.* Paper presented at the The ninth international workshop on presence (PRESENCE2006). Cleveland, OH.

Loomis, J. M. (1992). *Presence and distal attribution: Phenomenology, determinants, and assessment.* Paper presented at the Proceedings of the SPIE.

Lu, J., & Itti, L. (2006). Feature-based attention is not object-based. *Journal of Vision, 6*(6), 786a.

Luria, A. R. (1974). *Cognitive Development Its Cultural and Social Foundations* (M. Lopez-Morillas & L. Solotaroff, Trans.). Cambridge, MA.: Harvard University Press.

Maguire, K., & Connaughton, S. (2006). *Relational presence and distanced interdependent relationships.* Paper presented at the The ninth international workshop on presence (PRESENCE2006), Cleveland, OH.

Mania, K., & Chalmers, A. (2001). The effects of levels of immersion on memory and presence in virtual environments. *Cyberpsychology & Behaviour, 4*(2), 247-263.

Maxfield, L. (1997). Attention and semantic priming: a review of prime task effects. *Consciousness and cognition., 6*, 204-218.

Meehan, M. (2001). *Physiological reactions as an objective measure of presence.*, University of North Carolina at Chapel Hill.

Meehan, M., Insko, B., Whitton, M., & Brooks, F. P. (2002). *Physiological measures of presence in stressful virtual environments.* Paper presented at the SIGGRAPH 2002, San Antonio, USA.

Meehan, M., Razzaque, S., Whitton, M., & Brooks, F. P. (2003). *Effect of Latency on Presence in Stressful Virtual Environments.* Paper presented at the IEEE Virtual Reality 2003 Conference, Los Angeles, CA.

Mikropoulos, T. A., Tzimas, E., & Dimou, G. E. (2004). *Objective presence measures through electric brain activity.* Paper presented at the The seventh international workshop on presence (PRESENCE2004), Valencia, Spain.

Minsky, M. (1980). Telepresence. *Omni, 2*(44-52).

Monaco, J. (2000). *How to Read a Film: The World of Movies, Media, Multimedia: Language, History, Theory*. New York, NY: Oxford University Press.

Mori, M. (1970). Bukimi no tani (The uncanny valley). *Energy, 7*, 33-35.

Moye, J. (1997). Nonverbal Memory Assessment with Designs: Construct Validity and Clinical Utility. *Neuropsychology Review, 7*(4), 157-170.

Neisser, U. (1967). *Cognitive Psychology*. New York: Prentice Hall.

Nelson, R., Dickson, A., & Banos, J. (2000). An automated administration of Corsi's Block-tapping Test. *Perception and Motor Skills, 91*(2), 578-580.

Neter, J., Wasserman, W., & Whitmore, G. A. (1988). *Applied statistics*. Boston, MA.: Allyn and Bacon.

Nichols, S., Haldane, C., & Wilson, J. R. (2000). Measurement of presence and its consequences in virtual environments. *International Journal of Human Computer Studies, 52*, 471-491.

Nichols, S., & Stich, S. (2000). A cognitive theory of pretense. *Cognition, 74*(2), 115-147.

Nowak, K., Krcmar, M., & Farrar, K. M. (2006). *Examining the relationship between violent video games, presence, and aggression.* Paper presented at the The ninth internation workshop on presence (PRESENCE2006), Cleveland, OH.

Nunez, D. (2002). *A connectionist explanantion of presence in virtual environments.*, University of Cape Town, Cape Town, South Africa.

Nunez, D. (2004a, October 2004). *A constructionist cognitive model of presence.* Paper presented at the 7th International workshop on presence. (PRESENCE 2004), Valencia, Spain.

Nunez, D. (2004b). *Working memory and presence: Reconsidering the role of attention in presence.* Paper presented at the 7th International workshop on presence (PRESENCE 2004), Valencia, Spain.

Nunez, D., & Blake, E. H. (2001, November 2001). *Cognitive presence as a unified concept of virtual reality effectiveness.* Paper presented at the 1st International Conference on Computer Graphics, Virtual Reality, Visualization and Interaction in Africa (AFRIGRAPH 2001), Cape Town, South Africa.

Nunez, D., & Blake, E. H. (2003a, February 2003). *Conceptual priming as a determinant of presence in virtual environments.* Paper presented at the 2nd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (AFRIGRAPH 2003),, Cape Town, South Africa.

Nunez, D., & Blake, E. H. (2003b, February 2003). *A direct comparison of presence levels in text-based and graphics based virtual environments.* Paper presented at the 2nd International Conference on Computer Graphics, Virtual Reality, Visualization and Interaction in Africa (AFRIGRAPH 2003), Cape Town, South Africa.

Nunez, D., & Blake, E. H. (2003c). *A direct comparison of presence levels in text based and graphics-based virtual environments.* Paper presented at the 2nd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (AFRIGRAPH2003), Cape Town, South Africa.

Nunez, D., & Blake, E. H. (2003d). *The thematic baseline technique as a means of improving the sensitivity of presence self-report scales.* Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Nunez, D., & Blake, E. H. (2006). *Content knowledge and thematic inertia predict virtual presence.* Paper presented at the The ninth international workshop on presence (PRESENCE2006), Cleveland, OH.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Reviews Neuroscience, 4*, 95-102.

O'Keefe, J., & Reece, M. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus, 3*, 317-330.

Oberauer, K. (2002). to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 411-421.

Ohmi, M. (1998). *Sensation of self-motion induced by real-world stimuli.* Paper presented at the Selection and Integration of Visual Information: Proceedings of the International Workshop on Advances in Research on Visual Cognition,, Tsukuba, Japan.

Pinker, S. (1997). *How the mind works*. New York, NY: Norton.

Pinker, S. (2002). *The blank slate: The Modern Denial of Human Nature*. New York, NY: Viking Adult Press.

Plotkin, H. (1998). *Evolution in mind: An introduction to evolutionary psychology*. Cambridge, MA: Harvard University Press.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology., 32*, 3-25.

Posner, M. I., & DiGirolamo, G. J. (1998). Conflict, target detection and cognitive control. In R. Parasuraman (Ed.), *The Attentive Brain.* Cambridge, MA.: MIT Press.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition*. Hillsdale, NJ: Erlbaum.

Preston, L. (1998). *The use of virtual reality in the reduction of stress.* . Rhodes University, Grahamstown, South Africa.

Prothero, J., & Hoffman, H. G. (1995). *Widening the field-of-view increases the sense of presence in immersive virtrual environments.* (No. TR-95-2). Seattle, WA: Human Interface Technology Laboratory, University of Washington.

Prothero, J., Parker, D. E., Furness, T. A., & Wells, M. J. (1995). *Towards a robust, quantitative measure for presence.* Paper presented at the Conference on Experimental Analysis and Measurement of Situation Awareness.

Ravaja, N., Laarni, J., Kallinen, K., Saari, T., Salminen, M., Holopainen, J., & Järvinen, A. (2004). *Spatial presence and emotional responses to success ina video game: A psychophysiological study.* Paper presented at the The seventh international workshop on presence (PRESENCE2004), Valencia, Spain.

Reeve, B. & Nass, C. (1996). *The media equation: How people treat computers, television and new media like real people and places.* New York, NY.: Cambridge University Press

Relph, E. (1976). *Place and placeness.* London, UK.: Pion books.

Rettie, R. (2004). *Using Goffman's frameworks to explain presence and reality.* Paper presented at the The seventh international workshop on presence (PRESENCE2004), Valencia, Spain.

Rich, B. A., Schmajuk, M., Perez-Edgar, K. E., Pine, D. S., Fox, N. A., & Leibenluft, E. (2005). The Impact of Reward, Punishment, and Frustration on Attention in Pediatric Bipolar Disorder. *Biological Psychiatry, 58*, 532-539.

Riva, G., & Waterworth, J. A. (2003). Presence and the Self: a cognitive neuroscience approach. *Presence Connect, 3*(3).

Riva, G., Waterworth, J. A., & Waterworth, J. A. (2004). The layers of presence: A bio-cultural approach to understanding presence in natural and mediated environments. . *Cyberpsychology & Behaviour, 7*(4), 405-419.

Robillard, G. (2003). Anxiety and Presence during VR Immersion: A Comparative Study of the Reactions of Phobic and Non-phobic Participants in Therapeutic Virtual Environments Derived from Computer Games. *Cyberpsychology & Behaviour, 6*(5), 467-476.

Romano, D. M., Brna, P., & Self, J. A. (1998). *Collaborative decision-making and presence in shared dynamic virtual environments.* Paper presented at the Presence in Shared Virtual Environments.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behaviour research: Methods and data analysis.* New York, NY: McGraw Hill.

Rothbaum, B. O., Hodges, L. F., Alarcon, R., Ready, D., Shahar, F., Graap, K., Pair, J., Herbert, P., Gotz, D., Wills, B., & Baltzell, D. (1999). Virtual reality exposure therapy for PTSD Vietnam veterans: A case study. *Journal of Traumatic Stress, 12*(2), 263-271.

Rothbaum, B. O., Hodges, L. F., Ready, D., Graap, K., & Alarcon, R. (2001). Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder. *Journal of Clinical Psychiatry, 62*(8), 617-622.

Rumelhart, D. E., Hinton, G., & McClelland, J. L. (1986). A general framework for parallel distributed processing. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel Distributed Processing Vol 1: Explorations in the Microstructure of Cognition*: MIT Press.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro & W. E. Montage (Eds.), *Schooling and the Acquisition of Knowledge*: Lawrence Erlbaum Associates.

Ryan, M. L. (1999). *Cyberspace, Virtuality, and the Text. In Cyberspace Textuality: Computer Technology and Literary Theory*. Bloomington, ID.: Indiana University Press.

Sacau, A., Laarni, J., Ravaja, N., & Hartmann, T. (2005). *The impact of personality factors on the experience of spatial presence.* Paper presented at the The 8th international workshop on presence (PRESENCE2005), London, UK.

Sadowski, W., & Stanney, K. (2002). Measuring and managing presence in virtual environments. In K. Stanney (Ed.), *Handbook of virtual environments: Design, Implementation, and Applications.* (pp. 791-806). Mahwah, NJ: Lawrence Erlbaum Associates.

Sallnäs, E. L. (1999). *Presence in multimodal interfaces.* Paper presented at the 2nd International workshop on presence (PRESENCE1999), Essex, UK.

Sanchez-Vives, M., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience, 6*, 332-339.

Sander, I. D., & Galpin, V. C. (1994). A survey of attitudes to computing at the University of the Witwatersrand. In A. Adam, J. Emms, E. Green & J. Owens (Eds.), *IFIP Transactions A-57, Women, Work, and Computerization, Breaking Old Boundaries--Building New Forms*. Amsterdam: Elsevier Science.

Sas, C., & O'Hare, G. M. P. (2001). *The presence equation: An investigation into cognitive factors which underlie presence within non-immersive virtual environments.* Paper presented at the The 4th Annual International Workshop on Presence, Philadelphia, PA.

Sas, C., O'Hare, G. M. P., & Reilly, R. (2004). Presence and task performance: an approach in the light of cognitive style. *Cognition, Technology & Work, 6*(1), 53-56.

Schloerb, D. W. (1995). A quantitative measure of telepresence. *Presence: Teleoperators and Virtual Environments, 4*, 64-80.

Schneider, E. F., Lang, A., Shin, M., & Bradley, S. D. (2004). Death with a story: How story impacts emotional, motivational, and physiological responses to first-person shooter video games. *Human Communication Research, 30*(3), 361-375.

Schroeder, R., Steed, A., Axelsson, A., Heldal, I., Abelin, A., & Widestrom, J. (2001). Collabortating in networked spaces: As good as being there together? *Computers & Graphics, 25*, 781-788.

Schubert, T. W., Friedmann, F., & Regenbrecht, H. T. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments., 10*, 266-281.

Schubert, T. W., Regenbrecht, H. T., & Friedmann, F. (2002). Real and illusory interaction enhance presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 11*(4), 425-434.

Schuemie, M. J., van der Straaten, P., Krijn, M., & van der Mast, C. A. P. G. (2001). Research on presence in virtual reality: A survey. *Cyberpsychology & Behaviour, 4*, 182-210.

Schulte-Pelkum, J., Riecke, B. E., Caniard, F., & Bülthoff, H. (2005). Can auditory cues influence the visually induced self-motion illusion? *Perception, 34 (supplement)*, 82.

Sheridan, T. B. (1992a). Musings of telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments., 1*, 120-125.

Sheridan, T. B. (1992b). *Telerobotics, Automation, and Human Supervisory Control.* Cambridge, MA.: The MIT Press.

Shim, W., & Kim, G. J. (2001, May 2001). *Tuning the level of presence (LOP).* Paper presented at the 4th International worskshop on presence (PRESENCE2001), Philadelphia, PA.

Singer, M. J., & Witmer, B. G. (1999). On selecting the right yardstick. *Presence: Teleoperators and Virtual Environments, 8*(5), 566-573.

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence: Teleoperators and Virtual Environments, 8*(5), 560-565.

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

Slater, M. (2003a). A note on presence terminology. *Presence Connect, 3*.

Slater, M. (2003b). *Presence or pretence?* Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

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

Slater, M., Alberto, C., & Usoh, M. K. A. (1995a). *In the building or through the window? An experimental comparison of immersive and non-immersive walkthroughs.* Paper presented at the Virtual Reality Environments in Architecture, Leeds, UK.

Slater, M., Linakis, V., Usoh, M. K. A., & Kooper, R. (1996). *Immersion, Presence, and Performance in Virtual Environments: An Experiment using Tri-Dimensional Chess*. Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Slater, M., Pertaub, D., & Steed, A. (1999). Public speaking in virtual reality: Facing an audience of avatars. *IEEE Computer Graphics and Applications, 19*(2), 6-9.

Slater, M., Sadagic, M., Usoh, M. K. A., & Schroeder, R. (2000). Small-group behaviour in a virtual and real environment: A comparative study. *Presence: Teleoperators and Virtual Environments, 9*(1), 37-51.

Slater, M., & Steed, A. (2000). A virtual presence counter. *Presence: Teleoperators and Virtual Environments., 9*, 413-434.

Slater, M., & Usoh, M. K. A. (1993a). *The influence of a virtual body on presence in immersive virtual environments.* Paper presented at the 3rd Annual conference on virtual reality.

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

Slater, M., Usoh, M. K. A., & Steed, A. (1994). Depth of presence invirtual environments. . *Presence: Teleoperators and Virtual Environments, 3*, 130-144.

Slater, M., Usoh, M. K. A., & Steed, A. (1995b). Taking Steps: The Influence of a Walking Metaphor on Presence in Virtual Reality. *ACM Transactions on Computer Human Interaction, September 1995*.

Slater, M., Usoh, M. K. A., & Yiorgos, C. (1995c). *The influence of dynamic shadows on presence in immersive virtual environments.* Paper presented at the Selected papers of the Eurographics workshops on Virtual environments '95, Barcelona, Spain.

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.

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music, 19*, 110-120.

Snow, M. P. (1996). *Charting presence in virtual environments and its effects on performance.* Virginia Polytechnic and State University, Virginia.

Sperber, D., & Hirschfeld, L. (1999). Culture, Cognition, and Evolution. In R. A. Wilson & F. Keil (Eds.), *MIT Encyclopedia of the Cognitive Sciences* (pp. cxi-cxxxii). Cambridge, Mass.: MIT Press.

Squire, L. R. (1999). Memory, Human Neuropsychology. In R. A. Wilson & F. Keil (Eds.), *The MIT Encyclopedia of the cognitive sciences.* Cambridge, MA.: MIT Press.

Squire, L. R., Knowlton, B. J., & Musen, G. (1993). The structure and organization of memory. . *Annual Review of Psychology, 44*, 453-495.

Stanney, K., Mourant, R. R., & Kennedy, R. (1998). Human Factors Issues in Virtual Environments: A Review of the Literature. *Presence: Teleoperators and Virtual Environments, 7*(4), 327-351.

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

Stumpf, H., & Klieme, E. (1989). Sex-related differences in spatial ability: more evidence for convergence. *Perceptual and Motor Skills, 69*(3 Pt 1), 915-921.

Sutherland, I. (1965). *The ultimate display.* Paper presented at the International Federation of Information Processing Congress.

Timmins, L. R., & Lombard, M. (2003). *When the real seems mediated: Anti-presence.* Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Timmins, L. R., & Lombard, M. (2005). When "real" seems mediated: Inverse presence. *Presence: Teleoperators and Virtual Environments, 14*(4), 492-500.

Tong, F. (2002). Primary Visual Cortex and Visual Awareness. *Nature Reviews Neuroscience, 4*, 219.

Tooby, J., & Cosmides, L. (1990). On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation. *Journal of Personality, 58*, 17-67.

Towell, J., & Towell, E. (1997). Presence in text-based networked virtual environments or "MUDS". *Presence: Teleoperators and Virtual Environments, 6*, 590-595.

Treisman, A. M. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior, 3*, 449-459.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review, 76*, 282-299.

Treisman, A. M., Vieira, A., & Hayes, A. (1992). Automatic and preattentive processing. American Journal of Psychology. *American Journal of Psychology, 105*, 341-362.

Tromp, J., Bullock, A., Steed, A., & Frecon, E. (1998). Small group behaviour experiments in the COVEN project. *IEEE Computer Graphics and Applications, 1*, 53-63.

Tsur, R. (2002). Some Comments on the Lakoffean Conception of Spatial Metaphor. *Metaphor: A Multidisciplinary Approach. Special issue of Theoria et Historia Scientiarum., 6*(1), 245-267.

Tsur, R., & Benari, M. (2002). 'Composition of Place', Experiential Set, and the Meditative Poem (A Cognitive-Pragmatic Approach). *Pragmatics and cognition.*, 203-237.

Tulving, E. (1995). Organization of memory: Quo vadis? In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences.* (pp. 839–847). Cambridge, MA.: MIT Press.

Tulving, E., & Schachter, D. L. (1990). Priming and human memory systems. *Science, 247*, 301-306.

Turner, S. P., & Roth, P. A. (2003). *The Blackwell Guide to the Philosophy of the Social Sciences*. London: Blackwell.

Ueda, M. (1982). *Matsuo Bashō*. Tokyo, Japan: Kodansha International.

Ueda, M. (1996). *Modern Japanese Tanka*. Toyko, Japan: Kodansha International.

Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'Where' in the human brain. *Current opinion in Neurobiology, 4*, 157-165.

Usoh, M. K. A., Catena, E., Arman, S., & Slater, M. (2000). Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments, 9*(5), 497-503.

Usoh, M. K. A., Whitton, M., Bastos, R., Steed, A., Slater, M., & Brooks, F. P. (1999). *Walking > Walking-in-place > flying, in virtual environments.* Paper presented at the Computer Graphics Annual Conference (SIGGRAPH99), Los Angeles, CA.

Vandierendonck, A., Kemps, E., Fastame, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology, 95*(1), 57-79.

Vecera, S. P., & Rizzo, M. (2003). Spatial attention: normal processes and their breakdown. *Neurologic clinics, 21*(3), 575-607.

Vigotsky, L. (1978). *Mind in Society*. Cambridge, MA.: Harvard University Press.

Vinayagamoorthy, V., Brogni, A., Gillies, M., Slater, M., & Steed, A. (2004, 13-15 October 2004). *An investigation of presence responses across variations in visual realism.* Paper presented at the The seventh international workshop on presence (PRESENCE2004), Valencia, Spain.

Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, F., Böcking, S., Schramm, H., Gysbers, A., Hartmann, T., Klimmt, C., Laarni, J., Ravaja, N., Sacau, A., Baumgartner, T., & Jäncke, P. (2004). *MEC Spatial Presence Questionnaire (MEC-SPQ): Short documentation and instructions for application.*: Report to the European Community, Project Presence: MEC (IST-2001-37661).

Vorderer, P., Wirth, W., Saari, T., Gouveia, F. R., Biocca, F., Jäncke, F., Böcking, S., Hartmann, T., Klimmt, C., Schramm, H., Laarni, J., Ravaja, N., Gouveia, L. B., Rebeiro, N., Sacau, A., Baumgartner, T., & Jäncke, T. (2003). *Constructing Presence: Towards a two-level model of the formation of Spatial Presence.* : Unpublished report to the European Community, IST Program, Project "Presence: MEC" (IST-2001-37661).

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*(2), 250-270.

Waller, D., & Bachmann, E. R. (2006). The borderline of science: On the value of factor analysis for understanding presence. *Presence: Teleoperators and Virtual Environments, 15*(2), 235-244.

Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior research methods, instruments & computers, 35*(4), 550-564.

Waterworth, E. L., & Waterworth, J. A. (2003). The illusion of being creative. In G. Riva, F. Davide & W. A. IJsselsteijn (Eds.), *Being There: Concepts, effects and measurent of user presence in synthetic environments.* Amsterdam, The Netherlands: Ios Press.

Waterworth, J. A., & Waterworth, E. L. (2001). Focus, Sensus and Locus: The three dimensions of virtual experience. *Cyberpsychology & Behaviour, 4*(2), 203-213.

Waterworth, J. A., & Waterworth, E. L. (2003a). Being and time: Judged presence and duration as a function of media form. *Presence: Teleoperators and Virtual Environments, 12*(5), 495-511.

Waterworth, J. A., & Waterworth, E. L. (2003b). The core of presence: Presence as perceptual illusion. *Presence Connect, 3*(3).

Waterworth, J. A., Waterworth, E. L., & Westling, J. (2002). *Presence as performance: The mystique of digital participation.* Paper presented at the Presence 2002: The 5th International Workshop on Presence, Porto, Portugal.

Watt, R. J. (1988). *Visual processing: computational, psychophysical and cognitive research.* Hillsdale, NJ: Lawrence Erlbaum.

Welch, R. B. (1999). How Can We Determine if the Sense of Presence Affects Task Performance? *Presence: Teleoperators and Virtual Environments, 8*(5), 574-577.

Welch, R. B., Blackmon, 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, 5*(3), 263-273.

Wiederhold, B. K., Jang, D. M., Kaneda, M., Cabral, I., Lurie, Y., May, T., Kim, I. Y., WIederhold, M. D., & Kim, S. I. (2002). An investigation into physiological responses in virtual environments: An objective measurement of presence. In G. Riva & C. Galimberti (Eds.), *Towards CyberPsychology: Mind, Cognitions and Society in the Internet age.* Amsterdam, The Netherlands: IOS Press.

Wilson, J. R., Nichols, S., & Haldane, C. (1997). Presence and side effects: Complementary or contradictory? In M. Smith, G. Salvendy & R. Koubek (Eds.), *Design of computing systems: Social and ergonomic considerations*. Amsterdam: Elsevier.

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., Saari, T., Laarni, J., Ravaja, N., Gouveia, F. R., Biocca, F., Sacau, A., Jäncke, L., Baumgartner, T., & Jäncke, P. (2007). A Process Model of the Formation of Spatial Presence Experiences. *Media Psychology, 9*(4).

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005). The factor structure of the presence questionnaire. *Presence: Teleoperators and Virtual Environments, 14*, 298-312.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments., 7*, 78-89.

Youngblut, C. (2003). *Experience of presence in virtual environments* (No. D-2960): Institute for defense analyses.

Youngblut, C., & Hurie, O. (2003). *The relationship between presence and performance in virtual environments: Results of a VERTS study.* Paper presented at the IEEE Virtual Reality Conference, Los Angeles, CA.

Youngblut, C., & Perrin, B. T. (2002). *Investigating the relationship between presence and task performance in virtual environments.* Paper presented at the IMAGE 2002 Conference, Scottsdale, AZ.

Zeltzer, D. (1994). Autonomy, Interaction and Presence. *Presence: Teleoperators and Virtual Environments, 1*(1), 127-132.

Zimmons, P., & Panter, A. (2003). *The influence of rendering quality on presence and task performance in a virtual environment.* . Paper presented at the IEEE Virtual Reality 2003, Los Angeles, CA.