A CONNECTIONIST EXPLANATION OF PRESENCE IN VIRTUAL ENVIRONMENTS

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE,

FACULTY OF SCIENCE

AT THE UNIVERSITY OF CAPE TOWN

IN FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF PHILOSOPHY

By

David Nunez

February 2003

Supervised by

Edwin H. Blake

Abstract

Presence has various definitions, but can be understood as the sensation that a virtual environment is a real place, that the user is actually in the virtual environment rather than at the display terminal, or that the medium used to display the environment has disappeared leaving only the environment itself. We present an attempt to unite various presence approaches by reducing each to what we believe is a common basis – the psychology of behaviour selection and control – and re-conceptualizing presence in these terms by defining *cognitive presence* – the mental state where the VE rather than the real environment is acting as the basis for behaviour selection.

The bulk of this work represents the construction of a three-layer connectionist model to explain and predict this concept of cognitive presence. This model takes input from two major sources: the perceptual modalities of the user (bottom-up processes), and the mental state of the user (top-down processes). These two basic sources of input competitively spread activation to a central layer which competitively determines which behaviour script will be applied to regulate behaviour.

We demonstrate the ability of the model to cope with current notions of presence by using it to successfully predict two published findings: one (Hendrix & Barfield, 1995) showing that presence increases with an increase in the geometric field of view of the graphical display, and another (Salln?s, 1999), which demonstrates the positive relationship between presence and the stimulation of more than one sensory modality. Apart from this theoretical analysis, we also perform two experiments to test the central tenets of our model. The first experiment aimed to show that presence is affected by both perceptual inputs (bottom-up processes), conceptual inputs (top-down processes), and the interaction of these. We collected 103 observations from a 2x2 factorial design with stimulus quality (2 levels) and conceptual priming (2 levels) as independent variables, and as dependent variable we used three measures of presence (Slater, Usoh & Steed's scale (1995), Witmer & Singer's (1998) Presence Questionnaire and our own cognitive presence measure) for the dependent variable.

We found a significant main effect for stimulus quality and a significant interaction, which created a striking effect: priming the subject with material related in theme to the content of the VE increased the mean presence score for those viewing the high quality display, but decreased the mean of those viewing the low quality display. For those not primed with material related to the VE, no mean presence difference was discernible between those using high and low quality displays. The results from this study suggest that both top-down and bottom-up activation should be taken into account when explaining the causality of presence.

Our second study aimed to show that presence comes about as a result not of raw sensory information, but rather due to partly-processed perceptual information. To do this we created a simple three group comparative design, with 78 observations. Each one of the three groups viewed the same VE under three display conditions: high-quality graphical, low-quality graphical, and text-only. Using the model, we predicted that the text and low-quality graphics displays would produce the same presence levels, while the high-quality display would outperform them both. The results were mixed, with the Slater, Usoh & Steed scale showing the predicted pattern, but the Presence Questionnaire showing each condition producing a significantly different presence score (in the increasing order: text, low-quality graphics, high-quality graphics). We conclude from our studies that the model shows the correct basic structure, but that it requires some refinement with regards to its dealings with non-immersive displays. We examined the performance our presence measure, which was found to not perform satisfactorily. We conclude by proposing some points relevant to the methodology of presence research, and by suggesting some avenues for future expansion of our model.

Acknowledgements

Happy moments, sad moments, moments of exasperation, moments of jubilation – producing this dissertation has been a long, sometimes tiresome process, during which I experienced every emotion except one – loneliness. So many talented, enthusiastic (and even sarcastic) people helped me along the road, that it is difficult to know even where to begin expressing my gratitude.

Firstly, to Edwin Blake, my thesis supervisor, who bore the brunt of my overcomplicated experiments, bad writing, and incoherent ramblings. He did not always provide a solution to my problems, but he always maintained a positive attitude and far more than once it was his oddball take on a problem that provided me with the means to crack it. Why that man has so much faith in my abilities I'll never know; but if it were not for the confidence I drew from this, I would not be at this stage now.

Next, to Colin Tredoux and Gillian Fichilescu, who are my "research parents" having taught me almost everything I know about research methodology, statistics, and keeping an upbeat demeanor when the data are crumbling all around. Their guidance and simple advice during my research was of inestimable value.

Certainly this work could not have been produced without the aid of my lab mates, who have come and gone over the years. Their technical knowledge contributed much to my education in computing, and their optimism and support contributed much to my mental welfare. I hope I was able to return to them a fraction of the abundance they heaped on me. To this list add my parents, who although not too sure of the technical details, always made sure I worked when I had to, and stopped to play when I had to.

Finally, a word of thanks must go to two special contributors to this work: my brother, Fabian, who having completed his MSc, passed down all those little tricks and hints which are not printed on any postgraduate prospectus on earth; and to Cathryn Johns, who walked the long road with me, and understood and responded to my work and mental state better than anyone could dare ask for.

Contents

List of Tables

List of figures

Chapter 1

Introduction

"Professor [Thomas B.] Sheridan agreed that the development of a cognitive theory of presence would be a highly desirable goal. He suggested that 'pieces of it are lying around'"

Synopsis of general audience discussion, 1987 symposium of the NRC Committee on Human Factors (from Sheridan, Kruser, & Deutsch, 1987).

This dissertation deals with presence in virtual environments, and specifically with the problems of explaining and predicting presence. Presence has many definitions and it is difficult to define operationally (see chapter two for some of the most influential of these definitions), but it is not difficult to attain a "gut feeling" for this phenomenon. Presence has been defined as a feeling of "being there" in the virtual environment (Sheridan, 1992; Slater & Wilbur, 1995), or the sense that the medium you are watching disappears to leave only the scene you are watching (Lombard & Ditton, 1997). Others define presence as the desire to interact with computer controlled agents as if they were real people (Sheridan, 1996). Generally, presence is a feeling, derived from experiencing a virtual environment or other medium, that one is involved with the medium at more than face value, or that it is affecting one more than simply watching a scene would.

If presence is simply a feeling, then why is it, to paraphrase Sheridan, a highly desirable goal to develop a cognitive theory of presence? Again, the answer is complex, and the complexities are discussed fully in chapter two. The basis of the answer lies in the fact that presence is more than a simple feeling; the feeling is akin to the tip of an iceberg, whose submerged mass includes a vast number of potential psychological effects. These range from an improvement in the task being performed in the VE (Barfield, Zeltzer, Sheridan, & Slater, 1995) to a sense of companionship and social inclusion in a virtual group (Blake, Casanueva & Nunez, 2001). To maximize the benefits of presence, it is necessary to understand its causes and consequences. A theory is thus a necessity if one is to take advantage of these effects.

Many other researchers have created explanations of presence. However, none have, in our opinion, created an explanation of presence which takes enough advantage of recent findings in cognitive psychology. Most presence researchers agree that perception plays an important role in presence, although very few extend beyond perception to include the processes which control behaviour in the environment. We contend that the richness of psychological research into the link between perception, cognition and action (see Chapter three for a review) can be applied to the presence problem to create not only an explanation of presence, but also a means by which it could be predicted. We attempted to create such a theory by combining the pieces which Sheridan alerted us were 'lying around'. This dissertation contains our synthesis of findings from the presence research community with those of cognitive psychology, together with a preliminary empirical validation of our ideas.

1.1 Aims of the project

This project aims to produce a model which will both explain current presence findings in a unified way, and have the capacity to predict a subject's feeling of presence from a given set of initial conditions. To perform this task, we synthesized findings from the presence field with theories from cognitive psychology (see section Error! Reference source not found. below for more detail on the method used). The major theoretical aims of this dissertation are:

1. To attempt to unify the various strands of presence conceptualization by redefining them in terms of well-understood psychological processes which are responsible for ensuring environmentally appropriate behaviour. We call our unified concept *cognitive presence.*

2. To create a conceptual model to explain and predict cognitive presence. To do this, we make use of a connectionist architecture defined by McClelland & Rumelhart (1986). This model will include both top-down and bottom-up processes. This is, we believe, the first model of its kind in the presence field to give equal importance to these two fundamental perceptual processes.

3. To create a measure of cognitive presence. We wish to not require questionnaires or introspection. Preferably, it should be administered during the VE experience to minimize memory distortions. We call our measure the *Contents of Consciousness Inventory* (COCI).

4. To establish the validity of our theory by means of empirical investigation (see 1.2.2 below for more details of how this will be achieved).

Finally, we aim to investigate the psychometric properties our own presence measurement (the COCI), as we as of two published scales, namely the Slater, Usoh & Steed scale (1995), and the Presence Questionnaire of Witmer & Singer (1995).

1.2 The research

The method we used in this dissertation is unremarkable, and makes use of established empirical research practices from the fields of experimental psychology. We first engaged the theoretical goals of the project (outlined in 1.1 above), and then proceeded to empirically validate those ideas. Although our research aimed to support our theoretical model, we found one unexpected and powerful effect; namely, we found that creating an expectation in our users of the theme of the VE they were about to visit (by means of *conceptual priming*), lead to a differential effect on presence for users viewing high and low quality displays.

1.2.1 Theoretical development

We began the project by conducting a thorough analysis of current views and concepts of presence. Also, we investigated the views of cognitive psychology on the problem of maintaining behaviour appropriate to environmental conditions. Once this review was complete, we developed our concept of cognitive presence.

The next step was to create a method of measuring cognitive presence. We looked at the presence measurement literature, paying special attention to the criticisms of existing measures. We also reviewed the psychometrics literature, to determine the constraints and requirements of a psychological measure. From these elements, we created the concept of the Contents of Consciousness Inventory.

With these two goals achieved, we then turned to the central problem of creating an explanatory and predictive model of presence. We turned first to the presence modeling literature, to examine the successes and failures therein. Also, we looked at how cognitive psychologists model complex behaviour, and selected the connectionist architecture based on this review. We then proceeded to slowly create the model presented in Chapter five, by reviewing the literature at each stage to ensure that there was enough evidence to support each change we made to the model.

To ensure that the model we created was coherent with respect to the current body of presence literature, we selected two major published studies identifying causes of presence, and predicted their results with the model. The model was capable of predicting the actual findings (see Chapter 6 for the details of this process).

1.2.2 The empirical component

To test our theories empirically, we decided to identify a few of the foundation elements of our theories and test those. We reasoned that testing the foundations was preferable to testing the periphery, because a weakness in the foundations implies weakness in the periphery, but not vice-versa.

We identified the two major foundational components of our model to test:

1. The quality of the display will affect the degree of presence experienced by the user, and simultaneously, the user's mental state will affect how much they respond to the display.

2. Presence does not come about as a consequence of sensory information, but rather as a consequence of the mental processing of an environment.

Each of these points was investigated by means of an experiment. We designed each experiment and then used the model to predict the presence level in each case. We then conducted the experiments, and compared the empirical finding to the model's prediction

Experiment 1: This experiment investigated the degree to which display quality affects presence, and also the degree to which the user's mental state affects how they process the display. A secondary goal was to ascertain the validity and reliability of our presence questionnaire, the *contents of consciousness inventory* (COCI). Our hypotheses were that higher stimulus quality would improve the presence experience, and that priming (i.e. providing the subjects with information about the setting of the VE) would improve the presence experience. Details of this experiment are in Chapter eight. We used a 2x2 factorial design, with display quality and priming (manipulation of the user's mental state) as independent variables. The dependent variable was presence, as measured by Slater, Usoh & Steed's scale (1995), Witmer & Singer's Presence Questionnaire (1998), and our own COCI scale. We used two separate VEs, one representing a medieval monastery, the other a contemporary hospital. This was done to ensure generality; if the results were repeated in two such different settings, we would have more confidence in our findings.

The results from this experiment supported the first hypothesis (higher quality stimuli result in higher presence scores), but the second hypothesis was only partly satisfied. Specifically, we found a significant interaction between stimulus quality and priming. The effect of this is that priming lead to higher presence scores in the case of the subjects viewing the VE in a high quality display, but *decreased* the presence scores of those viewing the VE in a low quality display. This finding is quite surprising, particularly as the priming manipulation only related the *setting* of the VE (i.e. hospitals or monasteries) , and made no references to virtual environments or to display technology.

Experiment 2: This experiment tested the degree to which a text-based display is able to produce presence in subjects. Our hypotheses were that text displays would produce lower scores than text displays, and that priming would reduce the difference between text displays and graphical displays. We compared the presence scores collected on Slater, Usoh & Steed's scale (1995), Witmer & Singer's Presence Questionnaire (1998), and our own COCI scale of three groups of users. The first experienced a VE using a text-based display, and the other two groups viewed the same VE in the high quality and low quality graphical display conditions used in experiment 1. The details of this experiment are in chapter nine. We used a simple three group comparative design, with VE display as the independent variable, and presence as the dependent variable.

Our results showed that text displays produced lower presence scores that high quality graphical displays under all experimental conditions, but it was superior to low quality graphical displays under certain conditions. We surmise that this instability was due to artifacts of the questionnaires, particularly the PQ.

Apart from testing the basic tenets of our model, we also used the data collected in both experiments to conduct an analysis of the validity and reliability of our COCI scale, as directed in Anastasi & Urbina (1996). We also performed these analyses on the Slater, Usoh & Steed scale (1995), as well as Witmer & Singer's Presence Questionnaire (1998), to determine the quality of these scales.

1.3 Outline of this dissertation

Chapter 2: In Chapter 2, we present a discussion of presence, based on the literature available at the time of writing. We begin by discussing the various conceptualizations of presence which exist in the literature. We then present a detailed overview of presence measurement by considering five broad categories of presence measure, examining examples of each class, and providing a brief discussion of the relative strengths and weaknesses of each. The remainder of the chapter is devoted to an examination of presence models. We consider major classes of presence models, and provide a detailed examination and critical evaluation of six important models: the causal model of IJsselsteijn, de Ridder, Freeman & Avons (2000), Steuer's model (1992), Slater & Usoh's representation systems theory (1993), Thie & van Wijk's general presence theory (1998), the Immersion, Presence and Performance model (Bystrom, Barfield & Hendrix, 1999), and Schubert, Friedmann & Regenbrecht's path-analysis model (1999).

Chapter 3: Chapter 3 presents an introduction to the various psychological concepts which we make use of in the development of our model. Specifically, we focus on the theory created by cognitive psychologists to explain the selection and regulation of environmentally appropriate behaviour. We provide detailed discussions of the concepts of *schemata* (Rumelhart & Ortony, 1977) and of *scripts* (Schank & Abelson, 1977). The chapter closes with an introduction to connectionist architectures, with a detailed description of the workings of the interactive activation and competition architecture (McClelland & Rumelhart, 1986; Martindale, 1981), which we use for our presence model.

Chapter 4: This chapter introduces and discusses the problems created by the diversity of presence conceptualizations and measurement strategies which currently exist. We then critically consider Lombard & Ditton’s (1997) solution to this problem, and present an argument as to why that solution is not complete. The use of the psychology of behaviour selection (described in Chapter three) is suggested as another possible unifying component for presence, and we derive the notion of *cognitive presence* from this psychological research. We them consider the measurement of cognitive presence, and introduce the Contents of Consciousness Inventory (COCI) as a measuring instrument for this concept. The chapter closes with a critical discussion of our suggested solution.

Chapter 5: Chapter 5 defines our connectionist model of presence, and discusses how it relates to previous models (presented in Chapter two). After demonstrating the basic architecture, we provide a detailed description of each component, namely the conceptual layers, the action layer, and the perceptual analyzers. For each of these components, we discuss its derivation from the basic architectures defined by McClelland & Rumelhart (1986) and Martindale (1981). We also justify the structure and existence of each component in terms of psychological theory and findings from the presence field. We conclude the chapter with a discussion of how presence is expressed by the model.

Chapter 6: In this chapter, we present examples of how our model operates, by showing the reader how the model was used to successfully predict two published results in the presence field. We show how the model predicts the effect of manipulating the geometric field of view (Hendrix & Barfield, 1995), as well as the effect of using multimodal rather than unimodal display systems (Salln?s, 1999).

Chapter 7: This chapter proposes a strategy for evaluating the validity of our model. We describe four conditions which the model will be tested with, and use the model to predict a level f presence for each. The present the results which we expect, if the model is indeed valid.

Chapter 8: Chapter 8 describes the first experiment which we conducted as an implementation of the validation strategy described in Chapter seven. This experiment tested the relationship between priming and display quality, with respect to three measures of presence. The major finding is that priming affects the way that the display quality creates presence, and that, disregarding priming, display quality is a reliable cause of presence. We also found that the presence measures of Slater, Usoh & Steed (1995) correlates closely with the scale of Witmer & Singer (1998), and that both of these scales display good validity. We also found that our experimental COCI measure is not suitable for measuring presence.

Chapter 9: This chapter describes the second experiment we conducted, in line with the validation program outlined in Chapter seven. This experiment tested whether extremely low immersiveness displays (text-based) could induce presence, as predicted by our model. We found slightly mixed results, suggesting that text-only displays could induce presence as well as low-quality graphical displays, but not as high as high-quality displays.

Chapter 10: In Chapter ten we present a summary of the results collected in experiments 1 and 2. We conclude that the data provide some support for the validity of the model. The model was not able to predict the unique relationship between priming and presence. Also, the model predicted that, given the right conditions, non-immersive displays could bring about presence as effectively as immersive displays. The data collected in this regard was inconclusive, and so the question remains an open one. After considering these results, we discuss their implications for the model, and re-evaluate the model’s usefulness as a presence-predicting tool.

Chapter 11: This chapter presents some discussion of methodological issues raised during the conduct of this research, including an analysis of the psychometric properties of the presence measures used. We find that Witmer & Singer's scale presents high reliability, and adequate validity, although it is not suited for use with non-immersive displays. We also find that Slater, Usoh & Steed's (1995) scale shows unacceptably low reliability, but adequate validity. We then consider some of the implications raised by the experiments to the field of presence research, focusing on the role of priming, its possible existence as a confound variable, and how this may be overcome.

Chapter 12: This chapter concludes the dissertation by providing a summary of the obtained results, and the conclusions drawn. We close with some suggestions for future work.

Chapter 2

Presence: concepts, measures, models

The concept of presence in virtual environments is complex, and there exist many interpretations of it. It is impossible to enter into a substantial discussion of presence without first considering the various shades of meaning which compose the term *presence.* This diversity of meaning has lead to an equally diverse set of philosophies of presence measurement and explanation. This chapter considers the major schools of thought on presence conceptualization, and critically evaluates the various presence measurement methodologies. To close the chapter, we conduct a critical and detailed examination of a selection of models and other theoretical structures which have been proposed in the literature to explain and predict presence.

2.1 Immersion

A term which often appears in the presence literature is *immersion*. The term *presence* is complex and will be discussed in detail below, but the term *immersion*, although more simple, leads perhaps to more confusion, due to the fact that it has two quite different definitions. The first definition is psychological. In this sense, immersion refers to a feeling of being deeply involved in the virtual world, and entering it as if it were real (Coomans & Timmermanns, 1997, in Smith, Marsh, Duke & Wright, 1998), Witmer & Singer (1998) provide a more precise psychological definition for immersion. For them, immersion is "a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences" (Witmer & Singer, 1998, p. 227). They further explain that immersion in a virtual environment (VE) is a function of isolation from real-world stimuli, the perception of self-inclusion in the VE, natural modes of interaction with the VE, and the perception of self-motion through the VE.

It is clear from Witmer & Singer's definition that immersion (in the psychological sense) can be greatly affected by the hardware used to display the virtual environment. A head-mounted display (HMD), for example, provides more isolation from real world stimuli, and head tracking provides a natural mode of interaction. Similarly, a surround-sound system can provide a better sense of being surrounded by the VE than stereo speakers alone can. This notion is embodied in the second definition of immersion, which considers immersion in terms of display quality. Slater & Wilbur (1995) define *immersion* as a list of technologies which are necessary to induce presence. They agree with the notion that to be immersed is to be surrounded, but they argue that this can only be achieved with display technologies, such as head-mounted displays, which block out real world stimuli and provide virtual environment stimuli to the user's senses. It is thus possible to express the degree of immersion by simply stating the display technology used (Slater &. Wilbur, 1995). Schubert, Friedmann & Regenbrecht summarize this idea by stating, "immersion can be described objectively" (1999, p. 269). Witmer & Singer (1998) oppose this stance, arguing that immersion is a user experience, and although technology is the means by which immersion is achieved, immersion cannot be reduced to a list of the technology used to display it.

2.2 Telepresence and presence

Historically speaking, the concept of *presence* seems to have been derived from the term *telepresence*, which comes from the remote operation field. Akins, Minsky, Thiel & Kurtzman (1983, in Held & Durlach, 1992) reported that operators of high performance teleoperation systems often experienced the sense of being at the remote worksite rather than at the operator's terminal. Sheridan (1992) discusses this concept, and argues that telepresence arises from the mental representation that the teleoperator creates of the remote manipulator. Sheridan argues that if the remote manipulator presents a view of the remote site which matches that which the teleoperator would view were she at the site, and the lag between the operator's input and the feedback on those inputs were minimized, then telepresence would occur. Sheridan goes further, and states that the remote site need not be real; a virtual place with a virtual remote manipulator could also induce this sense of telepresence. To distinguish between telepresence felt for a real remote site and for a virtual site, Sheridan suggests referring to the latter as *virtual presence*.

2.3 Major schools of presence conceptualization

The rather straightforward concept of presence put forward by Sheridan (1992) has become more complex as more research has been done. Many researchers have since identified various forms of presence. Lombard & Ditton (1997) reviewed their contemporary conceptualizations of presence and created six categories of definitions, which they term *dimensions of presence*. Schuemie, van der Straaten, Krijn & van der Mast (2001) identify a further eleven uses of the term. Although there is a great deal of variety in these terms, we create three broad categories of presence conceptualizations for the purposes of this discussion (namely *social*, *personal* and *environmental* conceptions). Our categorization is similar to that used by IJsselsteijn, de Ridder, Freeman & Avons (2000), who summarize Lombard & Ditton's taxonomy into two broad categories: *physical* and *social*. Our first category, *social*, contains all concepts of presence which emphasize social interaction, communication, or the existence of entities in the VE other than the user. The second category, *personal*, includes conceptualizations which focus on individual users and their psychological states. The final category, *environment*, includes conceptualizations which emphasize the environment or task performed in the VE.

2.3.1 *Social* conceptualizations

Researchers working in the area of computer mediated communication and collaborative virtual environments are interested in the degree to which presence is capable of supporting intimate personal communications, collaborative work, and other social experiences (Lombard & Ditton, 1997). Such researchers consider users to be present in an environment when they are engaging in meaningful communication or other forms of social interaction. Several examples of this type of concept exist in the literature. Heeter (1992) proposes the term *social presence*, arguing that finding another person in the VE provides further evidence for the existence of the VE as a real place. She also argues for the importance of socially constructed realities for presence, which may arise as a consequence of having a group sharing a VE. These social realities will help to add meaning to the world, and further increase the sense that the VE is more than simple images (Heeter, 1995).

A term related to this idea is *co-presence*, which refers to the degree that a user feels that other agents in the VE represent real users (Durlach & Slater, 2000). Unlike social presence, co-presence is not thought to contribute to the personal varieties of presence (Slater, Sadagic, Usoh, & Shroeder, 2000), although there is evidence to suggest that the two are related (Blake, Casanueva & Nunez, 2001).

Slater, Usoh, Benford, Snowdon, Brown, Rodden, Smith, & Wilbur (1996) propose four aspects that are necessary to co-presence, and distinguish it from personal presence:

1. Personal presence must occur logically before co-presence.
2. A graphical representation of the user in the virtual world is necessary
3. The user must perceive the possibility of interaction or exchanging of information with the others in the VE; their static existence is not sufficient.
4. Graphical representations of the others in the VE. Slater *et al* (1996) suggest that the form of representation required may vary from user to user.

Lombard & Ditton (1997) propose the idea that a form of presence can also exist when the user acts as social actor inside the VE. They argue that when a VE contains other interactive entities, a present user will respond to them by behaving in accordance with the social rules which would apply to a similar situation in the real world. Cassell & Thorisson (1999) provide examples of how gaze direction and conversation pauses can be used by intelligent agents to encourage users to engage them with natural social modes. This type of agent in the VE can lead to another of Lombard & Ditton's dimensions, namely the environment performing as a social actor to the user. The interaction between agents and users in this way can, according to Delaney (1992, in Heeter, 1995), lead to a strong sense of social presence.

2.3.2 *Personal* conceptualizations

Sheridan's (1992) concept of presence is an example of the personal type. In Sheridan's view, presence occurs under two conditions. Firstly, the user must be viewing images of the VE from the perspective that the user would see if the VE were truly occupied. Secondly, the user must be able to interact with the VE in some way (even if only by moving through it), and must receive timely feedback for their actions. Once these conditions are met, then the subject will begin to feel if they are in the virtual world rather than in the room where the display equipment resides (Sheridan, 1992). This notion is quite similar to that expressed by Slater & Usoh (1993), who define presence as "the (suspension of dis-) belief that they are in a world other than where their real bodies are located" (Slater & Usoh, 1993, p. 222). Like Sheridan (1992), Slater & Usoh also see an important link between the hardware used (which they term *immersive properties*), and presence (Slater & Usoh, 1993). The basic notion of presence as "being there" is a common feature of many personal conceptualizations. Lombard & Ditton categorize definitions which emphasize this "being there" factor as *transportation* concepts (1997).

Heeter (1995) takes a slightly different line to Sheridan and Slater & Usoh. Rather than suggesting that presence spontaneously occurs when particular display criteria have been met, she argues that each of the features of the VE, display system or otherwise, provides the user with a piece of evidence about their existence in the VE. For instance, seeing your body (or self-representation) moving in a natural way in the VE presents convincing evidence that you are in the VE. However, there can also be evidence working to the contrary, such as poor rendering quality, and other features of the display system (Held & Durlach, 1992). On balance, however, the more positive evidence that exists, the more present the user will feel (Heeter, 1995). Zeltzer (1992) also works with a more active definition that those of Sheridan and Slater & Usoh. Zeltzer proposes that *interaction*, which he defines as "the degree of access to model parameters at runtime (i.e. the ability to define and modify states of a model with immediate response)" (Zeltzer, 1992, p. 127) is an important aspect of presence, and that interacting with an environment can contribute to the sense of presence. However, Zeltzer also recognizes the importance of the deterministic character of presence, and suggests a component simply called *presence*, which is similar to Slater & Usoh's use of the term. Unlike Slater & Usoh, however, Zeltzer does not consider this component alone to be enough for presence. He justifies this position by writing, "[w]e are immersed in a very high bandwidth stream of sensory input, organized by our perceiving systems, and out of this 'bath' of sensation emerges our sense of being in the world. This feeling is also engendered by our ability to affect the world through touch, gesture, voice, etc." (Zeltzer, 1992, p.128).

Schloerb (1995) argues against the use of transportation definitions (i.e. definitions which emphasize the sensation of being somewhere other than at the display apparatus), as he believes that making reference to user's feelings is imprecise; to replace it, he proposes the concept of *subjective presence*. According to Schloerb, a subject is said to be subjectively present when they cannot accurately identify whether they are in the real world, or in the virtual world. Subjective presence, says Schloerb, can exist in degrees, although it is not made clear how this would occur.

A final, significant, conceptualization to consider in our *personal* category is the one proposed by Lombard & Ditton (1997). They attempt to unify the various presence definitions, arguing that each shares a common core. They identify each concept as representing one or more aspects of an illusion of non-mediation. They define the *illusion of non-mediation* as the psychological state where a viewer fails to perceive the medium of communication, and responds as if the medium were not there (Lombard & Ditton, 1997). Lombard & Ditton suggest that there are four consequences for the user when this illusion of non-mediation occurs:

1. An invisible medium can provide stimuli which are more rich and meaningful. This includes both physical stimuli (of the VE) as well as social stimuli.

2. The dissolution of mediation means that the barrier between "inside the VE" and "outside the VE" disappears. Objects and persons previously on opposite sides of the barrier now share a common space

3. Actors in the medium will be perceived as being non-mediated, which in turn will encourage the user to respond to them with the social responses which are normally reserved for inhabitants of the real world

4. Because the barrier between the VE and the real world have disappeared, social cues produced by actors in the VE are more likely to be interpreted by the user as if they were real social cues coming from real persons.

2.3.3 *Environmental* conceptualizations

The final category of presence concepts groups those definitions which emphasize the role of the environment, or of the user's interaction with the environment. Two philosophical positions which have been applied to arrive at definitions of presence are those of Heidegger and Gibson (Sheridan, 1999). Each one of these positions expresses the relation between a person and their environment, and are thus suitable to understanding the relationship between users and virtual environments (Sheridan, 1999). Sheridan (1999) explains that Heidegger posits that it is quite difficult for a person in the normal course of everyday interaction to perceive the components of an environment. During normal interactions with the environment, tools are completely transparent to the user, and they are not thought of as independent objects (Zahoric & Jenison, 1998). However, when a breakdown in this pattern occurs, the disparate components of the situation become apparent. For instance, during hammering a nail into a wall, the user is not aware of the hammer; it is simply an extension of their will to have the nail in the wall. However, if the hammer should slip out of the hand, then the user would become aware of the hammer as a separate entity.

Gibson's view is similar (in Sheridan, 1999). For Gibson, perception exists primarily to acquire information from the environment about how action can be performed on it (to *perceive affordances*, in Gibson's terminology). For Gibson, affordances are not objects outside the user, but rather parts of the user's perception itself (Mantovani, Riva, 1999). This implies a strong link between perception and action (Gibson, 1979), because perception exists primarily to support appropriate action in the environment (Sheridan 1999).

These two environmental positions on perception have been incorporated into a presence conceptualization by Zahoric and Jenison (1998). Based on the positions of Heidegger & Gibson, Zahoric and Jenison arrive at an environmentally centered definition of presence: “[P]resence is tantamount to successfully supported action in the environment” (Zahoric & Jenison, 1998, p. 87). In this definition, the term *successful* means that when the user interacts with the VE, the VE responds in a way that seems logical and in concert with the user’s expectations. Zahoric and Jenison see this process as central, because, like Gibson, they see perception of an environment as existing only to allow meaningful action in that environment. In real environments, according to Heidegger and Gibson, action and perception exist in a feedback mechanism, which ensures that operation in the environment is always meaningful and directed (Zahoric & Jenison, 1998). Similarly, Zahoric and Jenison see presence as being possible only if the VE allows such a feedback mechanism to operate normally.

2.4 Measurement of presence

Many methods of measuring presence are currently available, but very few have gained widespread use. A few exceptions exist; for instance, Slater, Usoh & Steed’s questionnaire seems to have gained some acceptance, although it seems to have been abandoned recently by Slater in favour of physiological measures (Slater, 2002a; Slater 2002b). This variety in measurement practice doubtless arises from the youth of the subject, and inevitably leads to difficulties for researchers. Slater (1999) expresses this frustration succinctly: “at the end of the day, I use questionnaires because, for the time being, I do not know what else to do, and in order to construct predictive equations, concerned with how presence varies with other factors in groups of people, some method of quantification is necessary” (p. 564).

Presence researchers are currently trying a variety of techniques in search of an effective solution to the problem of measurement. Lombard & Ditton (2000) note that the lack of an established measure might be due to the wide variety of presence definitions which exist. This variety in definition, they note, leads to a variety in measurement philosophies. It is thus extremely difficult to list all the measures that are available, and complicating matters further is the problem that many measures are simply used in single studies, and are thus not validated. We propose thus to simply outline the major presence measurement methodologies, giving a notable measure as an example of each.

2.4.1 Introspective measures

This class measures presence by asking the subjects to report introspectively on their experience of presence. This is generally done retrospectively, usually directly after the exposure to the VE. There are exceptions to this, however. For instance, IJsselsteijn, de Ridder, Hamberg, Bouwhuis & Freeman (1998) and Freeman, Avons, Pearson & IJsselstiejn (1999) make use of a continuous variation method, where the user modifies their response during the actual VE experience. An early measurement method making use of introspection is found in Schloerb (1995). Schloerb suggests setting up a situation in which subjects wear a head mounted display (HMD), and their task is to state if the images they are seeing are of a real environment or of a virtual environment. The HMD worn by the user in each trial will either be displaying images of a virtual environment, or footage from cameras filming a real environment. Each subject makes many such judgments, and the measure of presence is derived from the ratio of correct identifications to incorrect identifications. This strategy attempts to define presence purely in terms of the system, however, as the measurement refers not to the subject, but rather to the system’s capacity to evoke presence in a user (Schloerb refers to this as *objective presence*). In any event, Schloerb did not implement his method, and there is thus no information on the method's validity or reliability.

Another well known example of the introspective method is the Slater, Usoh & Steed questionnaire (sometimes referred to as the SUS, for the authors' names). This scale has existed in several revisions. Examples of its use can be found in Slater & Usoh (1993), Slater *et al* (1994), Slater *et al* (1995) and Usoh *et al* (1999). The last major revision of this scale was presented in Slater *et al* (1995), and no significant changes have been published since.

The SUS consists of six items, each of which is an introspection task for the subject. The response format of each item is a semantic differential scale, allowing the subject an expression of degree. An example of an item of this scale is (semantic anchors in parentheses):

When you think back to the experience, do you think of the virtual environment more as images that you saw or more as somewhere that you visited?
(images that I saw) *(somewhere that I visited)*

This scale has not been submitted to a formal reliability test by its authors. The scale has also not been validated against an independent measure. The measure has, however, shown a fair degree of construct validity, as several studies have shown it to be sensitive to changes in the immersive properties of a VE, in line with the presence theory of its authors (e.g. Slater *et al*, 1994; Usoh *et al*, 1999, Slater & Usoh, 1993, and 2.5.3 below).

Another measure which uses the introspective method, is the Igroup Presence Questionnaire (Schubert, Friedmann & Regenbrecht, 1999). Note, however, that only some of the IPQ items are introspective. The IPQ consists of 13 items, which are aimed at measuring spatial presence (the sense that a VE is a real place). This scale also makes use of the semantic differential response format for quantifying subject responses. Several items of this questionnaire ask the subject to introspect on various aspects of the VE experience, for instance (semantic anchors in parentheses):

How real did the virtual world seem to you?
(about as real as an imagined word) *(indistinguishable from the real world)*

The original IPQ was created in German, and a great deal of reliability and validity research was conducted to determine its psychometric properties (Schubert, Friedmann & Regenbrecht, 2001). However, the English translation of the test has not been verified to be valid. Anastasi & Urbina (1996) warn that translating a test from one language to another can significantly alter its psychometric properties (particularly its reliability). Therefore, before the English version is widely adopted, further validation and reliability analyses should be undertaken.

The criticisms of introspection based measures are many and varied. Slater (1999) criticizes them due to their subjectivity. He argues that self-reports of experience are unsuitable for measuring presence, because the measurement becomes inextricably tied into personal aspects of the user. Even if two subjects had the same experience, he argues, it is unlikely that they would provide the same report. Slater's argument seems to extend itself to the reliability of these measure (i.e. the degree of noise it measures), but does not cover aspects of validity. Slater argues that other techniques (such as ethnography) could provide the same information as introspective measures, without the subjectivity penalty he perceives in introspective measures.

Introspection measures attempt to tap directly into the user's experience of presence. These measures assume that an individual has direct access to the information which makes up their experience. However, this might not be the case. A study by Nisbett & Wilson (1977) demonstrated the limits of introspection with a series of studies. In one study, subjects were asked to select, from a row of five identical pairs of stockings, the pair they thought was the best of the five. Once they had made their choice, they were asked to introspect as to why they chose that particular pair. Although most subjects chose they right-most pair, the explanations of their choice varied widely, with subject quoting the colour, quality of cloth, and various other reasons for their choice. Nisbett & Wilson argued that introspective reports do not function as memories of mental process, but rather they are a process of the subject constructing an explanation of their behaviour based on their personal theories of behaviour. These theories, argue Nisbett & Wilson, are held *a priori*, and are used by subjects to maintain a sense of meaningful action, even if their behaviour has no rational basis.

From the skepticism of Nisbett & Wilson we can infer that introspection is a less than ideal method of obtaining reports of a user's psychological state, particularly if the report occurs after the event in question, where memory artifacts can further obscure the experience (Eysenk & Keane, 1995). These limitations of introspection translate into a decrease in the construct validity of the scale, because such scales will, apart from measuring presence, also measure the na?ve hypotheses held by the subjects.

A recent development in introspective scales which partially overcomes Nisbett & Wilsons's (1977) objections is the Breaks in Presence method of Slater & Steed (2000). This method relies on the subject detecting when they experience a discontinuity in their VR experience, such as that which might be caused by an artifact in the rendering process (for example, a sudden change in frame rates, tripping over the cables of the display, etc). Slater and Steed argue that by counting the number of reports of these breaks in presence (BIPs), an estimate of presence can be obtained. To support the validity of their technique, they present data from one empirical study using 20 subjects. In this study, subjects were immersed in one of two VEs (High activity and Low activity), one of which demanded interaction with the VE and the other did not. Apart from performing the VE task, subjects were asked to vocalize when a BIP occurred; they were also administered the SUS presence scale (discussed above). Across both groups (High and Low activity together) Slater and Steed found a positive, significant correlation between the BIP measure and SUS scores ($r = 0.8$, $n = 20$). They did not evaluate the reliability of their method.

The BIPs method's most intriguing characteristic seems to be its ability to overcome the fundamental scaling problems inherent in introspection scales, by allowing the subject only to respond to a stimulus extreme (although this obviously restricts the measurement of the variance of the phenomenon and thus hides a certain amount of information). However, the BIPs method can be criticized on several points. Firstly, the evidence presented by Slater and Steed for the BIPs method is not particularly compelling, given the very small sample used. Their method of validation can also be questioned on the basis of their choice of only one other presence scale to compare with. When validating scales by means of the concurrent validation method, as Slater and Steed have done, it is always preferable to compare one's scale with as many other as possible, to reduce the likelihood that both scales are measuring the same third variable (Gregory, 1991; Cronbach, 1990). Their decision to use only a single measure to compare against seems odd, given the broad range of measures available in 2000, the year of that paper's publication.

The BIPs method can also be criticized, as Slater and Steed themselves admit (2000), on its theoretical grounds. They admit that there is not much evidence that presence occurs as a dichotomous phenomenon, although it is not unreasonable to suggest that breaks in presence are experienced as binary changes. A second criticism which they raise about their method, is the assumption that the requirement for the subjects to report BIPs will not affect their capacity to accurately report the BIPs. They point to findings from a study of the cognitive processing of ambiguous symbols (Girgus, Rock & Egatz, 1977 in Slater & Steed, 2000) which suggest that knowledge that a BIP might occur could increase the rate of BIP reporting, which would lead to under-estimates of presence. They counter this objection by arguing that subjects would only be conscious of the requirement to report BIPs after the BIP occurred; that is, the BIP itself would act as a recall cue to reporting the BIP. Slater (2002a) has recently argued that the BIPs technique could be improved by creating a physiological meaure of a BIP, such as by means of a polygraph. This technique has not yet been tested, so it remains to be seen if it will be effective (see 2.4.4 below for a discussion of physiological measures of presence).

2.4.2 Measures of cause

This class of measure is largely indirect, and based upon a causal model of presence. Although many of these measures still rely on self-report by the subjects, we do not classify them as introspective measures as they do not expect the user to report on subjective experiences such as feelings or impressions. Rather, these measures gauge the subjects' perceptions of various variables thought to cause presence, and so Nisbett & Wilsons's (1977) criticisms do not apply. From these perceptual self-reports, the measures infer presence which the subject must have felt, based on the causal model used by the scale. The best known of these measures is probably the Presence Questionnaire (PQ) of Witmer & Singer (1998). The PQ consists of approximately 30 items (the number varies depending on the revision of the test), each of which is of the semantic differential type. The items are grouped into four conceptual groups, termed factors (note that these groups were derived conceptually, not via a factor analytic method). Briefly, the four groups are (examples are drawn from Witmer & Singer, 1998):

1. *Control:* related to how well the subject was able to move or interact with the VE. For instance, *How natural was the mechanism that controlled movement through the environment?*

2. *Sensory:* related to environmental richness, consistency of information, perception of movement, etc. For example, *How closely were you able to examine objects?*

3. *Distraction:* related to the focusing of attention and awareness of the interface. An example: *How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?*

4. *Realism:* related to the realism of the scene and the meaningfulness of the experience, for instance, *How well could you identify sounds?*

One of the essential features of the PQ is its lack of any items which directly query the subject about the experience of presence. All four factors are measuring the perception (in a direct, Gibsonian sense) of display system features. This implies that the PQ measures *causes of presence* rather than presence itself. Presence causality has not yet been fully determined, so this means that some of the causes which the PQ measures must have been arrived at by theoretical rather than empirical means. This implies that the PQ must rely, in part at least, on some theoretical model of presence causality.

Slater (1999) points out that the PQ does not obtain a direct measure of the theorized causes of presence because it relies on the subject's perceptions of them. This, argues Slater, leads to a confound between the stimuli and personal aspects of the subject which may affect their perceptions. This argument seems plausible, but contains the assumption that stimuli are experienced directly by subjects. It assumes, for example, that a particular change in the display properties of the VR system will always be perceptible to subjects. Clearly, not all display changes will be perceptible to subjects. Asking subjects for their opinions of display features does not, as Slater seems to suggest, simply add noise to the measurement, but rather takes a measurement of a signal after it has passed through the filter of perception. It seems that the confound that Slater identifies might be insurmountable, because presence is a personal experience, and as such is grounded on perceptions rather than stimuli. Even if one accepts that the PQ is a measure of *perceptions of causes* rather than of causes themselves, the PQ still does not escape methodological criticism. Because the PQ relies on introspection of perceptions, the arguments of Nisbett & Wilson (1977) presented in 2.4.1 above apply here as well. However, as Eysenk & Keane (1995) point out, introspections of perceptions are usually far more accurate than introspections of experiences.

A different criticism (which can also be applied to Slater, Usoh & Steed's questionnaire) is that the PQ is designed to be administered after the VE experience. This intention is evident by the past-tense formulation of the items, as well as by the choice of a 30 item questionnaire, which would be cumbersome to use during a VE experience. This reliance on memory can lead to errors or distortions (Eysenk & Keane, 1995), and thus might affect the accuracy of the measurement. The relative importance given to each factor is of concern. The PQ simply sums all the responses, and thus each item contributes equally towards the total. However, each of the four factors do not provide an equal weight to the total, as there are not an equal number of items in each factor. For example, in version 2.0 of the PQ (which appeared in Witmer & Singer, 1998), the *Control* factor has 13 items (40% of all items), the *Sensory* factor 11 items (34% of all the items), the *Distraction* factor 6 items (19% of all items), and the *Realism* factor 7 (22% of all items). Furthermore, some items contribute to more than one scale. *Realism* shares items with *Control* (2 items) and with *Sensory* (1 item). *Distraction* shares one item with *Control*. Thus, the *Control* and *Sensory* factors contribute twice as much to the total compared to the *Distraction* and *Realism* factors. This uneven weight distribution is not discussed by the authors of the PQ to any degree of satisfaction, and without presenting any supporting evidence which suggests that control and sensory factors outweigh the others in determining presence.

The Igroup Presence Questionnaire (IPQ) of Schubert, Friedmann & Regenbrecht (1999) contains some items similar to those of Witer & Singer's PQ. These ask the subject to respond with their impressions of particular perceptions. For instance, subjects are asked to respond to *I was not aware of my real environment (fully disagree / fully agree)* and to *I felt like I was just perceiving pictures (fully disagree / fully agree)*. The discussions on the PQ presented above pertain to these items as well, due to their similarity to PQ items. However, the discussion should not be interpreted as applying to the scale as a whole, as it also includes a number of items requiring the subject to introspect on the presence experience directly; for those items, the arguments presented in 2.4.1 above should be applied.

2.4.3 Behavioural measures

This class includes those measures which attempt to detect presence by observing movement or posture changes in the user in response to the VE. The concept is simple: if the VE contains some structure or element which would trigger some movement or posture change if it were not virtual, and the subject responds to that element in the way that one would expect if that element were not virtual, then the subject is said to be experiencing presence.

Perhaps the earliest example of such a measure was the observation of reflex responses suggested by Held & Durlach (1987, in Sheridan, 1996), and later refined by Loomis (1992). This method assumes that particular behaviours are initiated with little volitional input, occurring mainly as responses to certain externally originating inputs. Held & Durlach (1987) provide an example. If an object (a ball, perhaps) is thrown towards a person, a reflex reaction will occur. This might vary from person to person (some may try to catch, and others may try to duck away), but the response occurs reliably in most people. Loomis (1992) argues that if such a response occurs towards a virtual object, this indicates a belief in the subject that the object is somehow real. Of course, because reflexes occur below the level of consciousness, the subject may still rationally object that the VE is not realistic or convincing; however, their reflex response to the thrown object, argues Loomis, shows a degree of presence.

This method is interesting because, unlike the introspective measures, the observation is not performed by the subject, and is carried out unobtrusively, thus reducing the influence of demand characteristics and other reactivity effects. Also, this measurement is done during the subject's VE experience, and so the undesirable artifacts of memory can be eliminated. This particular method was used with some success by Usoh, Arthur, Whitton, Bastos, Steed, Slater, & Brooks (1999). They included a room with a large hole in the floor in their VE in order to investigate a display manipulation. One of the presence assessment methods they used consisted of noting if subjects reacted to the virtual precipice in the same way they would to a real one (peering over the edge, making exclamations, etc.). This particular method is clearly not suitable for use in all settings. Many VEs simply do not include the elements which may lead to a reflex reaction (such as holes in the ground, flying balls, etc.). In fact, the inclusion of such elements in a VE could act to reduce the realism of the VE, which itself can be a threat to presence.

Less intrusive variations of this technique have been investigated. Sheridan (1996) suggests that spontaneous nonverbal social behaviours, such as automatic responses to greetings, could also be used to the same end. For example, Cassell, Torres, & Prevost (1999) argue that particular phrases and conversation structure allow participants in conversations to create a turn-taking order which allow conversations to continue without major interruptions. They argue that agents could be programmed to include these elements in their utterances to cue human participants and thus increase the realism of the interaction, and following Sheridan's suggestion we suggest that presence could be estimated by noting the degree to which users respond to such cues. Similarly, Cassell, & Thorisson (1999) demonstrate that gaze direction is used as a social cue that a person's attention has shifted. If a user is placed into an environment with agents displaying these cues, and responds appropriately to them, then this could be interpreted as a sign of presence.

Another technique for estimating presence based on behaviour is the observation of subjects' changes in posture during the VE experience. A subject's posture can change in response to haptic inputs (as occurs when one responds to the movements of the carriage when riding in a train), but it can also occur in response to visual inputs only (for instance, Lee & Aronson, 1974, in IJsselsteijn, Freeman, de Ridder, Avons & Pearson, 2000). Based on this premiss, Freeman, Avons, Meddis, Pearson & IJsselsteijn (2000) argue that these changes in posture could be used as an estimator of presence, even if the VR system does not include any type of haptic feedback. They subsequently discovered a noticeable difference in postural responses between subjects viewing stereo displays and those viewing non-stereo displays; this finding was later replicated by IJsselsteijn, Freeman, de Ridder, Avons & Pearson (2000). However, these measures of posture were not significantly related to introspective measures of presence in either study.

One possible interpretation of the data collected in these two studies is that the lack of relationship between the postural measure and the introspective measure suggests that the postural measures are not measuring presence. However, we feel that this interpretation is incorrect, as it assumes that introspective measure are valid measures of presence. Based on the discussion presented in 2.4.1 above, we feel that there is not enough evidence to make this claim. Another interpretation of the lack of relationship is that postural measures of presence and introspective measures of presence actually measure two slightly different concepts. Specifically, introspective measures quantify *presence as reported by the subject*, while postural measures quantify *presence as observed by the researcher.* Each of these concepts will include a particular bias, based on the motives of each individual. Also, each method requires a different form of interpretation to quantify. In the first instance, the subject must try to quantify a personal sensation; in the second instance the researcher must try to decide if particular movement was a change in posture or not. It is conceivable for these differences to explain the lack of relationship between these two classes of measure.

2.4.4 Physiological measures

Some researchers, such as Barfield & Weghorst (1993) suggest that presence can be objectively measured by means of physiological measurements. They suggest the use of electromyography, electrocardiography and ocular motion. Prothero, Parker Furness & Wells (1995) suggest that electroencephalography could also be used. More recently, Meehan et al (2002) made use of electrodermal response, electrocardiography and respirometry. The attractiveness of these measures seems to lie with their objectivity. Because the measurements are of physical phenomena, it is possible to achieve extremely high levels of reliability (Gregory, 1991). However, as Anastasi & Urbina (1996) note, increasing the reliability of a measure does not guarantee a corresponding increase in its validity. They argue that the most important indicator of a measurement's quality is its validity (the degree to which the scale measures the construct it aims to measure). An increase in the reliability of a scale, they argue, simply shows the scale's ability to discriminate the signal from the surrounding noise. Reliability, however, says nothing about whether the scale is measuring the correct construct or not. Prothero, Parker, Furness & Wells (1995) add to weight to this argument by stating that no physiological correlates of presence have been yet identified, and thus not enough evidence exists to support the hypothesis that presence can be estimated by observing the physiology of the subject.

Lombard & Ditton (2000) agree with Prothero *et al* (1995) in principle, but argue for an exception. They propose that presence in a special class of virtual environments (those which are expected to create detectable physiological changes in users) could be evaluated by this method. An example of such an environment was used by Meehan, Insko, Whitton, & Brooks (2002), which presented users with a sudden precipitous drop. In such environments, where one can expect to create a reaction of fear, then physiological changes, such an increase in heart rate and a decrease in skin resistance, can be reasonably expected. Apart from such rather special cases, however, physiological methods offer little promise. Sheridan (1992) summarizes the arguments against physiological measures by stating,"[i]t is natural to seek an objective measure or criterion that can be used to say that telepresence or virtual presence have been achieved. However, telepresence (or virtual presence) is a subjective sensation, much like mental workload, and it is a mental model – it is not so amenable to objective physiological definition and measurement" (p. 209).

2.4.5 Comparisons to other mental states

A new trend in presence measurement is to estimate presence by means of indirect psychological measurement. This technique is quite similar to the behavioural measurement method, but does not rely on visible behaviours to indicate presence. This method is based on the simple idea that being in any environment leads to particular psychological effects, and measuring the degree to which the virtual environment is eliciting such effects can be used as a measure of presence.

Based on this notion, simulator sickness is sometimes used as an inverse measure of presence. Simulator sickness occurs as a consequence of the mismatch between information presented to the visual system and to the vestibular system (Kennedy, Hettinger & Lilienthal, 1988), which can often arise from experiencing a poorly implemented simulation (Pausch, Crea & Conway, 1992). Biocca (1992) describes simulator sickness as a syndrome containing a host of symptoms which are quite similar to those experienced in motion sickness. Its use in measuring presence is twofold. Firstly, simulator sickness implies that the simulation is not complete (Pausch *et al*, 1992), and thus is unlikely to lead to presence. Secondly, subjects experiencing simulator sickness will experience difficulty in maintaining attention focused on the virtual environment (Witmer & Singer, 1998). This is expected to lead to a decrease in presence (Alison, Harris, Jenkin, Jasiobedzka, & Zacher, 2001). Based on these arguments some researchers argue that simulator sickness could be used to estimate presence. Witmer & Singer, for instance, used simulator sickness as a sign that subjects were not present during the validation of their Presence Questionnaire (Witmer & Singer, 1998).

The use of simulator sickness as a presence measure is attractive for various reasons. Firstly, there exists a large body of research on the relationship between simulator variables, physiology and simulator sickness. Secondly, a reliable and valid measuring instrument, the Simulator Sickness Questionnaire, already exists to measure it (Kennedy, Lane, Berbaum & Lilienthal, 1993). Finally, there is some evidence (for example Miller, Sharkey, Graham & McCauley, 1993) to suggest that simulator sickness can be evaluated physiologically, which again raises hope for an objective means for estimating presence.

Simulator sickness, however, may not be a suitable method for measuring presence. Firstly, there exists very little information on the relationship between simulator sickness and presence, which means that it is lacking validation as a presence measure. Secondly, even if simulator sickness is found to impede presence reliably, that would only mean that one of the necessary conditions for presence to occur would be the lack of simulator sickness. To assert that simulator sickness can be used as an inverse measure of presence is to imply that presence and simulator sickness exist in a perfect negative relationship. However, it is quite simple to imagine situations where simulator sickness and presence would both be zero, or conversely, both be maximized. Consider, for example, a VE which simulates sitting on a park bench. There will be no mismatch between the visual and vestibular systems (and hence no simulator sickness), as the subject is sitting in both the laboratory and the VE. However, if the scene is executed poorly, very little presence is likely to result. In this case, both simulator sickness and presence would be at zero (or at least, very low). A second example illustrates the opposite phenomenon. Consider a VE of an airplane doing aerobatics. In this situation, there is likely to be simulator sickness, as the vestibular and visual systems will be giving quite different messages to the user. However, the VE is one in which it might be appropriate to feel motion sickness (whose symptoms are similar to those of simulator sickness), so in this case simulator sickness might actually contribute to the presence experience.

Another different technique which makes use of comparisons is that suggested by Waller, Hunt & Knapp (1998). They argue for investigating *fidelity* – the extent to which the user's interactions with the VE are indistinguishable from the participant's observations of and interactions with a real environment. This is a highly useful metric, as it provides a measure not only of how much the VE is impacting on the user's mental state, but also allows a direct assessment of the degree to which the VE experience might impact performance on a particular task (for example, learning). Mania & Chalmers (2001) implemented this technique by comparing memory performance across four conditions: real world, HMD display, desktop display and audio only display. Once assigned to a condition, subjects experienced a lecture-style presentation in the environment. Mania & Chalmers then tested subjects' memory for the content of the presentation, as well as for the space in which it was experienced. They found that presence (as measured by introspective scales) did not correlate with recall performance, but they did find that the type of memory for the spaces differed significantly across conditions, suggesting that manipulating the depth of immersion leads to a detectable difference in environment recall.

It would be incorrect to dismiss Mania & Chalmer's (2001) presence assessment method by arguing that they failed to find a correlation between their measure and other measures of presence. The fallacy of such a statement would be threefold; firstly, such a criticism assumes that there exists a universally accepted and operationalized concept *presence*, and as has been discussed in 2.3 above, this is currently not the case. Secondly, the statement assumes that current presence measures are valid and established, which as this discussion on measurement attests to, is currently not the case. Finally, even if a universal concept of presence existed, with a good measure of it, the statement assumes that presence correlates with all effects that a VE might have on a user. Indeed, these three arguments can be applied to any broad criticism of measuring presence by means of measuring variables other than presence itself. Conversely however, it is also incorrect to assume that the answer to measuring presence lies in reducing presence to another variable.

2.5 Models of Presence

Many reviews of presence research, such as the recent paper by Schuemie, van der Straaten, Krijn & van der Mast (2001) contain a section discussing causes of presence. In this dissertation, we opted to not include such a section, for several reasons. Firstly, we choose to avoid a review of causes as this dissertation concerns itself with presence modeling and presence prediction – it would be far more fruitful for our purposes to consider causes as elements of models rather than as free standing entities. Secondly, we feel that speaking about *causes of presence* gives the false impression that presence is a single, unified concept which is widely accepted and understood. The discussions in 2.3 and 2.4 above, we believe, suggest that this is not currently the case. The current state of presence measurement, in particular, makes it quite difficult to separate a finding from the method which was used to obtain it. For example, studies by Slater and colleages (e.g., Slater, Usoh, & Steed,1994 or Slater & Usoh, 1993) measure presence by means of introspective reports. However, studies by IJsselsteijn and colleagues (e.g. IJsselsteijn, Freeman, de Ridder, Avons, & Pearson, 2000 or IJsselsteijn, de Ridder, Freeman & Avons, 2000) measure presence by means of posture observation. These two measures have not been shown to be related (Schuemie & van der Mast, 1999), and therefore, we believe, findings on the causes of each should not presented in the same review.

For the purposes of this section, we define a *presence model* as an explanatory structure whose purpose is to *predict* the occurrence or level of presence a VE user will experience under a particular set of circumstances. This prediction may occur quantitatively or qualitatively. This definition excludes conceptualizations or definitions of presence, unless these make specific reference to possible causes of presence. Hughes (1994) provides a set of four desiderata for positivist models. Briefly, these are:

1. Models should be based (partly, at least) on empirical evidence. This is necessary to ensure that the model has a basis upon observation, rather than on conjecture.

2. Models should not simply be *post-hoc* explanations of previous findings. Hughes (1994) notes that it is always possible to create a structure to explain a set of past occurrences, making such structures correct *a priori*.

3. Models should be predictive, not postdictive, structures. Eysenk & Keane (1995) argue that the validity of a model exists not only in its ability to account for past observations alone, but equally so in its ability to predict future occurrences.

4. The evidence provided for the existence of a model should include instances of the model's ability to remain consistent to findings made before its inception, as well as instances of validated predictions made using the model.

Note that the points raised by Hughes also act to create the distinction between a *model* and a *definition*. Although definitions are sometimes based on empirical findings, definitions have no predictive power. Definitions are structures whose purpose is to identify, rather than predict.

2.5.1 Classes of presence model

Based upon the criteria outlined in 2.5 above, we identify four classes of presence model. Each class is defined in terms of its structure, its use of empirical findings, and whether it predicts the presence experience quantitatively or qualitatively. Note that this classification is based solely on a survey of the literature, and is not meant to provide a complete taxonomy of all possible models. The classification we propose contains four categories, listed below. Discussions and examples of each category are presented in 2.5.2, 2.5.3, 2.5.4 and 2.5.5 below.

1. *Unstructured causal models:* These models make use of a flat structure, with short causal chains. Typically, these models assume that all named causes work directly on presence, without mediation by other factors. These models tend to be strongly based on empirical findings, and they are usually capable of qualitative predictions only. These models often limit themselves to listing variables which have been empirically established as causes of presence.

2. *Structured causal models:* These models represent an added level of structural complexity above that of the unstructured causal models. They do this by identifying a structure which establishes relationships between the identified causes. These models also tend to be qualitative in their predictions, although they are not as reliant on empirical findings for determining their structure, which is usually extracted from theory.

3. *Structured conceptual models:* These models are usually the most complex, positing various levels of hidden structure (such as feedback loops) between the identified causes of presence, and presence itself. These models usually rely on empirical findings to identify input and output variables, but the structure connecting them is derived theoretically. These models tend to predict presence qualitatively.

4. *Regression models:* These models aim to quantitatively predict presence by use of single or multiple regression, factor analyses, path analyses or some other statistical method. These models are usually strictly empirically driven, without a strong theoretical basis.

2.5.2 Unstructured causal models

These models explain presence by relating it to a set of cause variables. Typically, no structure is present, as each cause variable is thought to act directly on presence. No postulation of mediator or moderator variables typically exists in these models. These models are not generally referred to as models or theories in the literature, and are typically not expressed explicitly. This type of model has appeared in the literature under the names of “factors affecting presence” (Barfield & Hendrix, 1995a), “factors underlying presence” (Witmer & Singer, 1998) and “determinants of presence” (Loomis, 1992; IJsselsteijn, de Ridder, Freeman & Avons, 2000). Although the discourse used by these authors seems to imply their desire to not have these structures regarded as models, the way in which they are used suggests that they are indeed models, because they are often used to make predictions or generate hypotheses about presence. These models are created, it seems, by an accumulation of published empirical findings. As each new finding is published, the model grows.

IJsselsteijn, de Ridder, Freeman & Avons’ Model

An example of such a model is present in IJsselsteijn, de Ridder, Freeman & Avons (2000). In a section entitled “determinants of presence” they provide a list of four classes of variable which they argue have an effect on presence. Their list is comprised of variables identified through “theoretical analyses” (p. 396), which have not been tested in the laboratory, and variables identified by empirical procedures. Their classification includes:

1. *The extent and fidelity of sensory information.* This is defined as the amount of useful and relevant information presented in a consistent way to the sensory modalities of the user. IJsselsteijn *et al* provide the examples of monocular and binocular cues of spatial layout, display resolution and spatial audio.

2. *The match between sensors and the display*. This is defined as the mapping between a subject's action in the VE, and the consequences of those actions in the VE. The example provided is that of the link between head-tracking and camera movement.

3. *Content factors*. This is defined as a broad category which encompasses the objects and physical structure of the VE. This includes the subject's representation in the VE (the avatar), as well as the interactions which are possible between the subject and the VE. IJsselsteijn *et al* suggest that the task carried out in the environment, as well as the meaningfulness of its content also belong to this category.

4. *User characteristics*. These factors include the subject's various faculties, including perceptual, cognitive and motor ability. It also includes prior experience with VR, expectations, and "willingness to suspend disbelief" (p. 3961). Huang and Alessi (1999 in IJsselsteijn *et al*, 2000) also include states of psychological disorder, such as depression and phobia, into this category.

5. *Negative cues*. This is not included as a category by IJsselsteijn *et al*, but we include it, as it is explained in some depth in their treatise. Unlike the other four categories above, negative cues do not contribute to presence, but rather detract from it. This includes all manner of variables, including a lack of user attention to the VE, eye strain, outside interruptions and system lag.

Apart from creating these five categories, IJsselsteijn *et al* group categories 1 and 2 together as being *media factors*, which will contribute towards the illusion of non-mediation. These media factors are the factors in the model most grounded on empirical findings, while the others are based mostly on conjecture.

We place IJsselsteijn *et al*'s model into the *unstructured causal models* category for several reasons. Firstly, the model is clearly an attempt not only at explanation, but also at prediction. The list of variables is created in an attempt to show the relative importance of some factors over others (an explanatory goal), while at the same time, they are used in an attempt to predict presence levels in particular situations. For example, considering the large number of media related variables listed, it is clear that presence is somehow connected to the fidelity of the stimuli presented to the user (this is an explanation). At the same time, by examining the negative cues, it is possible to predict that presence will be less if the VE experience occurs in a noisy room than if it occurred in a quiet one (this is a prediction).

IJsselsteijn *et al*'s model predicts presence qualitatively. It provides a mechanism for predicting presence in a relative sense (for instance, condition set *A* will lead to more presence that condition set *B*), but provides no way of determining the quantity of presence which will be felt. Furthermore, all the factors of this model are weighted equally, as no concrete indication is given of which factors may be of more or less consequence.

This model presents several strengths and weaknesses, most of which are common to all models in its class. The most significant weakness is its treatment of previous findings. This is a problem because the model is perceived by researchers not as a theoretical structure, but rather as a statement of proven facts (by virtue of their having been discovered in the laboratory). Because these statements are "known to be true" there is no perceived need to verify them. This is, however, not always the case. Even if the results used to build these models are statistically significant, they are often derived from modest samples (in the region of 40 subjects), and are not often replicated. This leaves no way of verifying that the relationship found was not a product of the particular circumstances in which it was discovered.

A second weakness is the model's lack of cognizance of the relative strengths of the factors which compose it. Although the studies upon which such models are based almost always report the statistics necessary to compute effect sizes, this is not often done. More importance seems to be given to the fact that a particular variable has a statistically significant relationship to presence, than to the strength of that relationship. Although population effect sizes are difficult to compute accurately from small samples, it should be possible to create a table of "major factors" and "minor factors" from the information available in the literature. Creating such a list would be difficult, as it is rare to find effect sizes published in presence research.

A third, important weakness of these models is their disregard for structure. As can be seen in IJsselsteijn *et al*'s model, all variables are considered as having a direct relationship with presence, and none are identified as mediators. This lack of structure may arise from the list-like nature of the models, which often aim to be understood simply as summaries of empirical findings. However, it is important to note that this lack of structure often simply reflects the research on which the models are based, as not many published studies exist whose aim was to identify mediators rather than main effects. It is entirely possible that when mediator variables begin to be identified, they will begin to be incorporated into these models.

This class of model, however also presents a notable strength, namely its reliance on empirically derived information. Although some untested factors are sometimes built into these models (as is the case in IJsselsteijn *et al*'s model presented above), these variables are usually identified as such, and far more importance is granted to those derived in the laboratory. This is a highly desirable property in any model, as it ensures that the information used in the model's construction truly exists.

2.5.3 Structured causal models

Structured causal models show more complexity than unstructured causal models. Like unstructured model, structured model still focus on basic causes of presence, but they go further by defining structures to relate the causes to one another. Unlike their unstructured counterparts, structured causal models are largely perceived as being theories in their own right. A well-known example of such a model is that presented by Steuer (1992). Steuer's model is similar to that of IJsselsteijn *et al*, in that it clusters causal variables into categories; however, Steuer imposes a structure onto these variables, thus creating a more complex predictive structure.

Steuer's Model

Steuer uses mainly variables which have been empirically tested, as well as variables which arise from theoretical analysis. Steuer's model is only capable of predicting presence at a qualitative level, as it does not incorporate any scalars of either cause or effect. Steuer identifies two major dimensions of VR display technology which contribute to presence. These are:

1. *Vividness.* This refers to the ability of the technology to present a rich set of stimuli to the user. According to Steuer, many factors contribute to the vividness of a display, but two are most significant. The first of these is *breadth*, which refers to the capacity of the system to stimulate a variety of sensory modalities. The second factor is *depth*, which refers to the amount of data which is carried to each of the stimulated modalities. When perceiving real environments, notes Steuer, the senses are almost always stimulated to maximum depth and breadth. However, virtual environments present information which is limited in both depth and breadth.

2. *Interactivity.* This dimension is also noted as being a feature of the display system. Steuer defines it as the extent to which users of a VE can modify the form or content of that VE. Interactivity includes several aspects which have been identified by various other researchers as being important to presence. These include Sheridan's notion of autonomy (Sheridan, 1992), and Zeltzer's notion of interactivity (Zeltzer, 1992). Steuer elaborates on the concept of interactivity by positing some of the factors from which it is composed. These include *speed,*

range and *mapping*. Speed refers to the rate at which the environment is able to update in response to user input. Range is the number of avenues of change which are open to the user at any time, and mapping refers to the ability of a system to respond to user input in a natural and predictable manner. Schuemie & van der Mast (1999) review the results of various studies investigating the importance of these interactivity components. They report evidence that suggests that presence is decreased noticeably if the speed of the system drops below a certain threshold value (which exists in the range of 15Hz to 20Hz). The evidence for range and mapping however, is far less clear. Schuemie & van der Mast conclude that the evidence does suggests neither a simple nor a monotonic relationship between presence and range. The data collected on the relationship between mapping and presence is similarly contradictory.

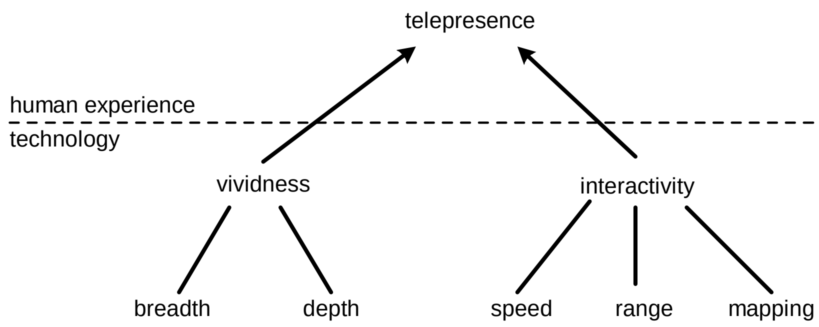


Figure 2-1: Steuer's model of presence (from Steuer, 1992). The bottom layer represents the five basic determinants of presence. These are grouped into to major factors ('vividness' and 'interactivity') of technological contributors to presence.

Apart from these two dimensions which he describes in detail, Steuer also includes a vaguely specified and unnamed *user psychology* dimension (this is our term for it; Steuer does not include it as a separate entity). This dimension includes all aspects of the presence experience which are contributed by the user's own internal state. A final dimension of Steuer's model is the *situational dimension*. This is constituted of transient factors in the environment which modify the situation; for instance, a sudden loud noise which might affect the presence experience.

Apart from these factors, Steuer also identifies the "willing suspension of disbelief" (p. 91) as an important ingredient of presence. Steuer believes that presence is not attainable unless the subject is willing to treat the mediated images as non-mediated to, at least to some extent. This process, according to Steuer, is mostly volitional, but there can be some unconscious forces (such as social context influences) at play as well.

Steuer's model does not stop at the careful classification of variables into dimensions. It also includes a structure to organize them. The structure, presented in Figure 2-1, shows a hierarchical arrangement of dimensions and components. The dotted horizontal line represents the interface between the generated display and the user's psychology. If the vividness and interactivity of the display are to have an effect on the user (that is, make their influence felt as presence), the user must first suspend disbelief. Thus, the dotted line represents not only the interface, but also a portcullis which the user must raise before the possibility of presence can exist.

Steuer's model presents several notable strengths and weaknesses. Its major strengths are increased explicatory power over the simple unstructured models, provided by virtue of its structure. It is possible to predict, for instance, that there are more opportunities for a system designer to increase the interactivity of a system than to increase its vividness. Also, more specific explanations are possible thanks to the structure. For example, it is possible to explain why an increase in display rate increases presence (it is an increase in *speed* and thus *interactivity*).

However, the structure of the model may turn out to be only a fa?ade of complexity. Note that all dimensions in the model carry the same weight. For all its apparent complexity, the model is actually a flat causal model. Because *breadth* and *depth* affect *vividness* equally, the *vividness* node only acts to categorize the terms. If the vividness node were removed, the explicatory power of the model would be unaffected (this would not be the case however, if the *vividness* node applied some transformation to these inputs). Similarly, the *interactivity* node can be removed with no detrimental effect to the model. These two dimensions, it seems, act only as semantic categorizations, rather than functional dimensions. The model can thus be reduced to an unstructured causal model, with almost no negative impact on its powers of prediction.

The major contribution of Steuer's model is its inclusion of user psychology as an important determinant of presence. Although Steuer does not include it as a separate construct in his model, he makes a strong case for its importance by making it pervasive in most other parts of the model. However, Steuer presents no empirical evidence to support his claim. This is unfortunate, as the *willing suspension of disbelief* is an important feature of his model. Steuer himself admits that "[t]his process is of great interest in the context of all kinds of mediated experience" (p. 91). Its lack of empirical basis, however, raises the question of whether users of VEs engage in this type of behaviour at all. Sadly, the lack of an empirical basis is a weakness which can be applied generally to Steuer's model. It relies substantially on the theoretical statements made by others, which are, in turn, not based on empirical data. This weakens the model, as predictions made with it will not be based upon observed patterns, but rather on expected patterns. While considering this weakness, as serious as it is, it is important to remember that this model was created in the earliest days of presence research, before much empirical research was available to build models from.

2.5.4 Structured conceptual models

This class includes those models which rely on higher level abstract concepts and complex structure to make predictions. They differ from structured causal models mostly in the specificity of the inputs to the model. Structured conceptual models use very high level concepts as inputs, whereas structured causal models have very specific inputs. For instance, Thie & van Wijk's model (a structured conceptual model described below) has an input node *signals* which encodes all of that encoded by Steuer's *breadth, depth, speed, range* and *mapping* nodes. Another difference exists in the complexity of the structure. Structured conceptual models typically have multiple layers of complexity, and often incorporate feedback loops. Structured causal models, on the other hand, are usually far simpler in structure.

There are several models of this class in the literature, representing a great theoretical variety. We will discuss three models which we feel are most significant in theoretical terms. We present them in chronological order of publication. They are Slater & Usoh's representation systems theory (1993), Thie & van Wijk's general presence theory (1998) and Bystrom, Barfield & Hendrix's Immersion, Presence and Performance model (1999).

Slater & Usoh's representation systems theory

Slater & Usoh build their model upon concepts derived from neurolinguistic programming therapy (NLP). This psychological basis provides only one half of their model; the other half is provided by empirical research into the causes of presence. The model posits a split between *external factors* and *internal factors*. External factors are those which are related to the display of the virtual environment. These factors include (Slater & Usoh, 1993):

1. High fidelity information presented to the sensory modalities of the user in such a way as to suggest non-mediation to the user. This factor includes Steuer's notion of "vividness" (see 2.5.3 above).

2. Presentation of the environment to the user in a way that is consistent across sensory modalities.

3. The user should be able to interact with the environment, and the environment should respond to the user by, for example, containing agents which interact with the user.

4. The self-representation of the user in the virtual environment should be similar to the user's real body, and should respond in an accurate way.

5. The effects of interacting with the environment should be simple enough for the user to learn over time. This is similar to IJsselsteijn *et al*'s notion of mapping (see 2.5.2 above).

Slater & Usoh note that this list is not complete; other display properties also exist as external factors. It is worthwhile noting that the *external factors* component of this model is remarkably similar to Steuer's *technological* layer, although Slater & Usoh's conception is far more abstract, encoding many more concepts into each factor. For instance, the *high fidelity* category of Slater & Usoh's model includes Steuer's concept of vividness as well as its components (i.e. breadth and depth). By making use of this increased abstraction, Slater & Usoh are able to include more causes (both established and speculated on) into their model without overloading it.

Any presence model can clearly benefit from the inclusion of psychological knowledge, as presence is rooted in personal experience. Making use of established research in psychology to aid in the construction of a presence model is thus a useful strategy. Slater & Usoh exploit this strategy by deriving a substantial portion of their model (the *internal factors*) from the NLP literature. These internal factors consist of three *representation systems* which act to process perceptual inputs, as well as other, more abstract thought structures. The three representation systems are:

1. *Visual representation system (V):* This system encodes all visual information. Apart from this basic sensory function, this system also contains self-constructed visual images, and visual memories

2. *Auditory representation system (A):* This system processes sound stimuli to produce auditory perceptions. This system is also used to produce the internal dialog (the silent conversations we have with ourselves) and holds auditory memories.

3. *Kinesthetic representation system (K):* This system is responsible for producing perceptions of touch, bodily motion and pose. It encodes memories for these sensations, as well as emotions.

These three systems are not all equally important in their contribution to experience. Each person has one system which will tend to be preferred over the others. This dominant system tends to contribute more when operating in an environment, although all three systems contribute to thought at every moment.

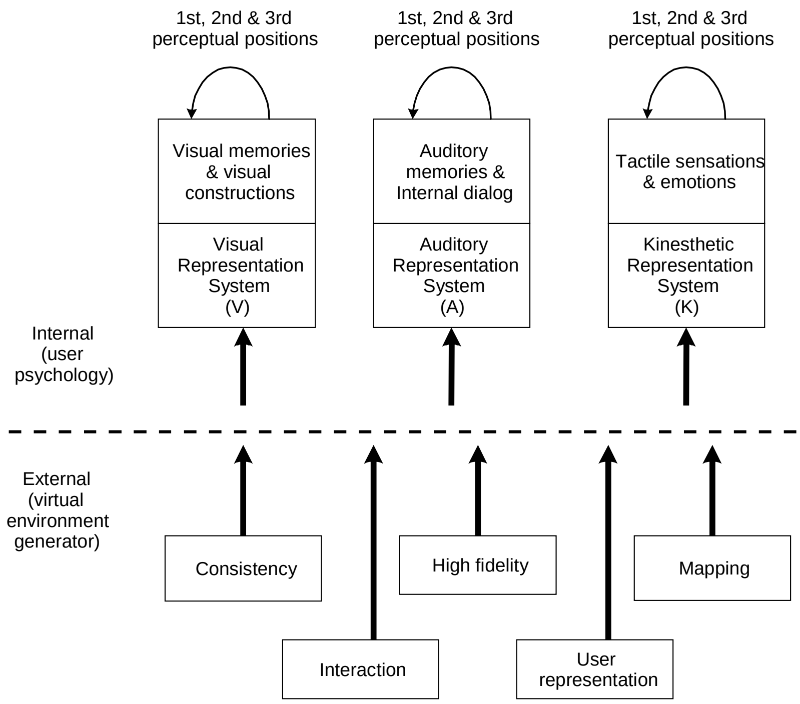


Figure 2-2: Slater & Usoh's representation system theory. The external factors stimulate the three internal representational systems, which lead to perceptions and presence (if the 1st perceptual position is dominant). In each individual, one of the three representational systems is dominant over the others; stimulation of this dominant system will lead to more intense sensations of presence.

Apart from the three systems described above, NLP theory also contains the notion of three *perceptual positions*, closely related to the three voices of normal speech. The first of these is referred to as the *first position*, which involves recollection of events as if they were occurring to the subject. The *second position* recalls events from the perspective of another person, and in the *third position*, the events are remembered from a passive, non-personal perspective. The important difference between a perceptual position and a voice in speech is that unlike voices (which only apply to speech), perceptual positions apply to any of the three representation systems. For instance, recall of a scene could occur from the perspective which the subject experienced it (first position), or from the perspective of someone else in that scene (second position), or from a "floating camera" perspective, not attached to any person (third position).

Slater & Usoh combine the internal and external factors in a structure reminiscent of Steuer's (Figure 2-2). In this model, equal weight is given to the external factors, but not to the internal; consistent with NLP theory, they posit that at any time, one of the representation systems will be more active than the others. An interesting feature of this model is that it has a distributed rather than explicit representation of presence. In this model, presence occurs when the subject is operating in the first perceptual position in one or more representation system. As one of these representations systems is more active than the others, presence will be experienced more strongly if the VE display equipment supports that modality well. For instance, a subject in a VE with compelling visuals who has a dominant visual representation system will experience more presence than one, in the same VE, who has a dominant auditory representation system.

Slater & Usoh published, together with their model, two sets of results aimed at verifying the model. They present some evidence to support the model; they found that the sense of presence correlated with the use of first position predicates in subject's descriptions of their experiences. This supports the distributed expression of presence in the model. However, the samples used for these studies were unsuitably small – 20 subjects in the first experiment and six in the second. These samples seem even smaller when one considers that they were used to construct a multiple regression equation with four predictors; this technique usually requires larger samples to ensure adequate statistical power (Howell, 2002). The findings should thus not be regarded as definitive or final.

Apart from this lack of empirical evidence, there are some other criticisms which can be leveled at the model. Although the strategy of using previous research in a related field as a basis for a presence model is a commendable one, Slater & Usoh's choice of NLP is less so. NLP is a controversial technique, which itself lacks adequate empirical verification. Slater & Usoh's admit to this, stating "the evidence for the NLP approach is largely anecdotal, practitioners implicitly accepting its correctness by virtue of their success in clinical (and other) applications. In fact, NLP has had a critical response from the academic and counseling community" (p. 223). Their selection of NLP is thus quite puzzling, considering the vast range of psychological theories examining perception and its link to behaviour that exist (see Chapter three for a broad outline of some of these theories). Slater & Usoh defend their choice by saying, "[w]e are not taking a standpoint here on the validity of the NLP claims, but rather using the model as a basis for the formation of testable hypotheses of interest in our IVE research." (p. 224) However, this defense is not valid, because a model is a predictive structure; if any of its parts or structure is false, its predictive quality will suffer as a result. In building a model, each component must be carefully selected, because its addition might lead to a decrease in the validity or accuracy of the model. For instance, if it should emerge that NLP's precepts are false, Slater & Usoh's model will likely not survive the consequences, as the internal factors would evaporate, leaving it as a variation of IJsselsteijn *et al*'s model.

Thie & van Wijk's general presence theory

Thie & van Wijk (1998) use a communications theory perspective for their model, and so it carries a distinct social flavour. Thie & van Wijk's central stated goal is to create a model which can encompass task performance. The structure of this model is presented in Figure 2-3 below. A strong component of this theory is the idea by Loomis (1992) that operation on a VE leads to the creation of a mental model of it. The concept of a mental model, which can be traced back to Craik (1943), has received considerable attention, via Loomis, in the presence research world. The concept is simple. As a person operates in an environment, they learn a series of cause-effect links between their actions and the consequent changes in the environment; this is referred to as a mental model. As this mental model grows, it becomes possible for the subject to use it to create expectations about the consequences of action in the environment, and consequently, complex planned actions become possible. In VE research, the mental model which a user has of the VE has importance for this same reason – if a correct model of the virtual environment has been formed, complex behaviours (such as the completion of some task) become possible in that VE.

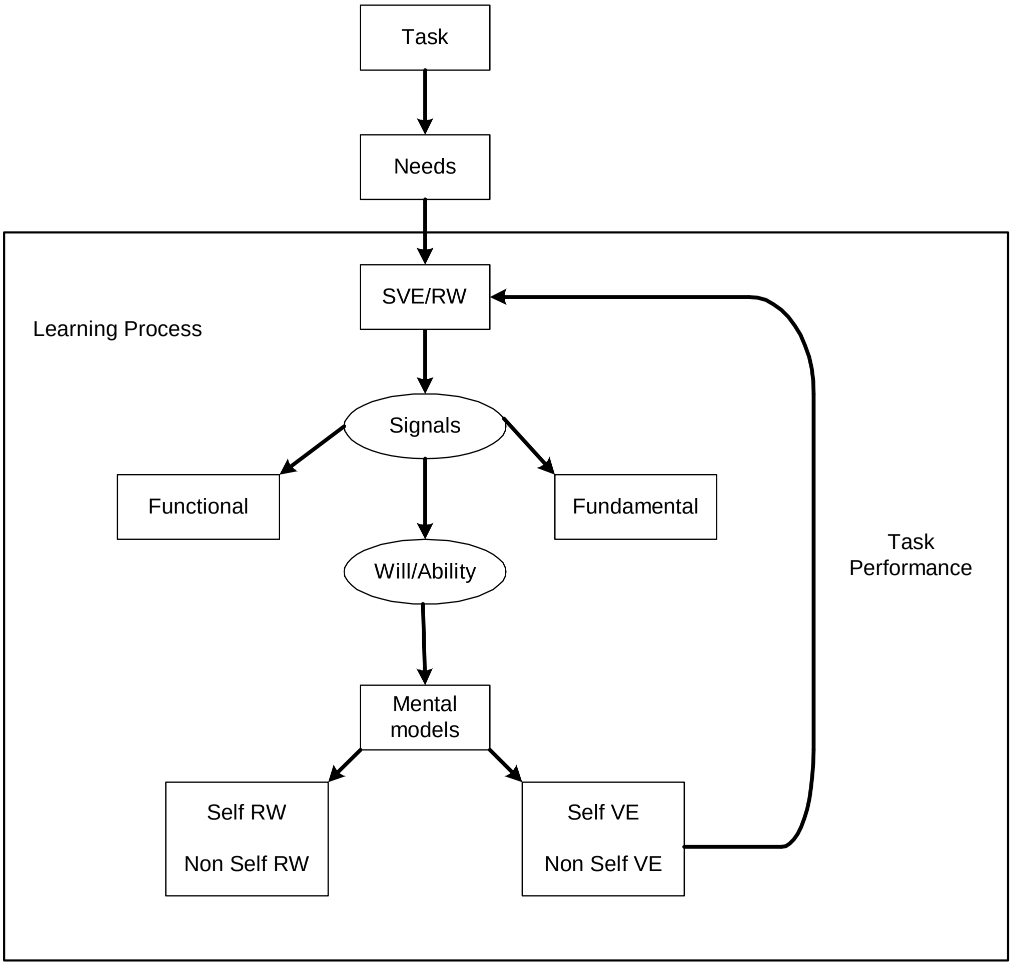


Figure 2-3: Thie and van Wijk's model (after Thie and van Wijk, 1998). Presence occurs when the mental model of the virtual world (Self VE/ Non-Self VE) becomes highly stimulated. This leads to learning, which feeds back and leads to increased performance in the VE.

Thie and van Wijk contend that any time a user is in a VE, two mental models are competing for activation: the model of the real world (RW), and that of the shared virtual world (SVE). When the subject is experiencing a high level of activation of the virtual mental model (labeled "non-self VE" in Figure 2-3), presence occurs. The activation of the SVE model rather than of the RW model, depends mostly on two factors: the willingness of the subject to become present, and the degree to which the subject feels there are others present in the VE.

1. *Willingness to become present:* Thie and van Wijk argue that a user's psychology contributes to their presence in two ways; one conscious, and the other unconscious. The conscious component they define as the user's willingness to accept signals from the VE. The unconscious element seems to be linked to the focusing of attention, as it is defined as the ability to block out signals from the RW while accepting signals from the VE. According to Thie and van Wijk, increasing these two components should lead to an increase in presence.

2. *Perception of others in the VE:* This trait, which is referred to as social virtual presence by Thie and van Wijk, has also been called co-presence by others (see Blake, Casaneuva & Nunez, 2000, for a brief review of this concept). Social virtual presence is regarded by Thie and van Wijk as a measure of the quality of communication in a VE. Thie and van Wijk argue that if communication with others in the VE is effective, then this will lead to an increase in the user's sensitivity to the social context, and a degree of de-individuation. This can lead to a change in the user's model of themselves in the VE, and consequently to a change in the degree of presence felt.

According to this model, presence occurs when the subject has a correct, developed mental model of both the virtual environment, and of themselves as occupants of that environment. Presence is further enhanced by the existence of other subjects in the environment, each of which is also present. Task performance is postulated to be important, due to its relationship to the mental model of the VE. As the mental model of the VE develops, argue Thie and van Wijk, task performance will improve due to the improvement in quality of information about the environment which a correct mental model brings. Thus, improved task performance can lead to an increase in presence. Also, when subjects are present, task performance will improve because the active mental model which is relevant (i.e. the mental model of the VE) will be relevant to the task. This will only occur, however, if the VE is well suited to supporting the task. It is also seems possible that the activation of a particular VE mental model could lessen performance on a task, if that environment is not developed for supporting that task, although Thie & van Wijk do not discuss this eventuality.

Thie and van Wijk present a modest set of results to support their model. They report a significant but moderate relationship between measures of social virtual presence and presence, as predicted by the model. However, they fail to find a relationship between measures of susceptibility-for-presence (measuring the *will/ability* component of the model), and presence. They argue that the erratic pattern of findings point not to a flaw in the model, but rather to a deficit in the measures. They support this argument by quoting unacceptably low Cronbach's alpha scores for the scales of both presence and susceptibility-for-presence. Sadly, because Thie and van Wijk did not make use of any other presence measures in their study, we are left with no clear evidence of their model's validity.

Thie and van Wijk's work is interesting in that it attempts to include ecological elements into the model, in the form of task performance. This explicit attempt to account for the "usefulness" of presence (insofar as it could aid task performance) is in some respects an improvement over Slater & Usoh's model, which remains cryptic as to the possible applications of presence. Another improvement of this model over that of Slater & Usoh, is that this model acknowledges the role of learning and experience in presence, whereas Slater & Usoh model the user as a static entity.

However, this model shows several serious flaws. Firstly, it expresses presence in a way that is too narrow to allow the application of the model to scenarios beyond those for which it was specifically developed. For instance, it is difficult to predict the effect presence might have on the emotions of the user (a prediction which is possible with Slater & Usoh's model). Also, although a distinction is made between conscious and unconscious factors of the user's willingness to experience presence, very little is explained as to how these operate as separate entities. This raises the question of whether this division adds any value to the model, or simply acts to violate Occam's razor. Finally, Thie & van Wijk's model suffers from a lack of empirical support. Not only were the findings of the studies which aimed at validating the model not convincing with regard to the model's correctness, but they also did not address some of the model's central contentions, such as the role played by the acquisition of the VE mental model, or the impact of task performance on this process.

The immersion, presence and performance (IPP) model (Bystrom, Barfield & Hendrix, 1999)

Bystrom, Barfield & Hendrix (1999) propose a model which also aims to clarify how task demands and task performance relate to presence and display technology. This model is itself an amalgam of previous models (mostly of the structured causal variety), and as such is built upon a solid foundation. The IPP model is a purely qualitative one, and according to its authors, exists more as a research tool than an engineering one (Bystrom, Barfield & Hendrix, 1999).

The model begins with two basic assumptions. The first of these is that presence can only occur as a consequence of immersion (an axiom proposed by Slater & Wilbur, 1995). The second assumption is that there exists a relationship between presence and performance (this axiom is proposed by the authors of the model). The model is presented in Figure 2-4 below.

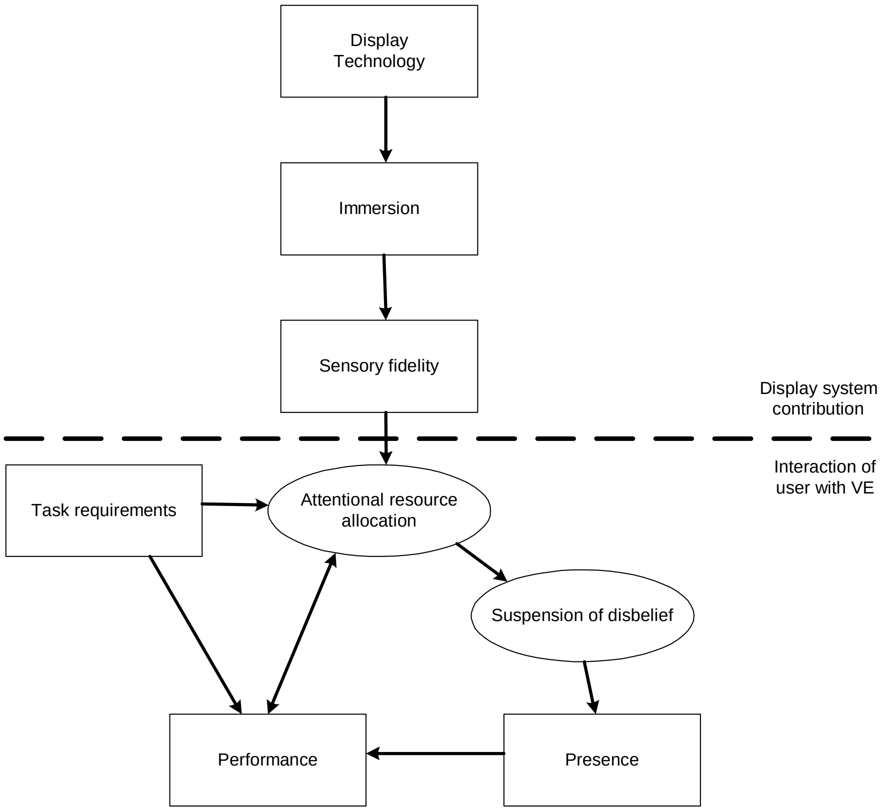


Figure 2-4: The IPP model (after Bystrom, Barfield & Hendrix, 1999)

As was the case with the models of Steuer and of Slater and Usoh, the model begins with a basic split between display technology and user psychology. The IPP model posits a simple causal relationship between the display system and its effect on presence, similar to that proposed by Steuer (1992) and Slater and Usoh (1993); namely, high fidelity immersive displays will lead to presence (although there is no explicit requirement of fidelity expressed, so we assume that the IPP allows for lowers levels of presence can arise from the use of non-immersive displays).

The first node in the user component of the model is the allocation of attention to process the information arising from the displays. Bystrom *et al* require that attention be allocated to the display before presence can occur. This requirement is quite similar to Thie and van Wijk's requirement that the VE mental model be more active than the RW mental model. Once attention has been focused, there must be some degree of suspension of disbelief before presence can occur. Unfortunately, Bystrom *et al* are vague as to whether this suspension must be done willingly by the subject, or whether it occurs automatically as a function of attentional resource allocation. This point is crucial, as according to the structure in Figure 2-4, presence can only come about as a function of this suspension of disbelief.

An interesting feature of this model is how it explains the possible relationship between presence and task performance. Unlike Thie and van Wijk's model, which proposes a rather crude direct link between presence and performance, the IPP theorizes that presence is affected by a bidirectional link to the allocation of attentional resources. As can be seen from Figure 2-4, performance can affect the allocation of attentional resources, which in turn can affect presence. However, presence itself can also affect task performance, and so it is possible for a positive feedback loop to be established to maximize both presence and task performance. However, the model does not suggest that task performance is a cause of presence or vice-versa; their intensities are both by-products of the allocation of attention.

Bystrom *et al* do not present any empirical data to support their model. Instead they argue that the model's primary purpose is not quantitative predictions, but rather the stimulation of research. It seems to us, however, that a model does not have to be able to make quantitative predictions for it to be a useful engineering or research tool. If this is to be done, however, then it is necessary to verify the IPP model by testing its essential features.

The IPP model exhibits some problems beyond those of its lack of validation. The most significant is the vague definition and handling of the contribution of the user's psychology to the presence experience. The allocation of attention is well explained, but the *suspension of disbelief* is not. This problem is shared with the model of Thie and van Wijk, which explains the unconscious contribution of the user (quite similar to the allocation of attentional resources concept used by Bystrom *et al*), but fails to satisfactorily define or explain the conscious element. The only detail Bystrom *et al* give is that "[i]f the participant allocates sufficient attentional resources to the virtual environment, and if there is a sufficient degree of sensory fidelity, the participant *may* 'suspend disbelief' and view the virtual environment as a real place" (p. 234, our emphasis). It would seem that if the suspension of disbelief is a prerequisite to presence in the IPP (as is implied by the structure in Figure 2-4), then more information about this phenomenon is required to completely understand this model.

2.5.5 Regression models

This final class of model differs from the other three in that it specifically aims to predict presence quantitatively. To do this, these models rely on various statistical techniques, notably regression and associated techniques. Due to this statistical basis, a successful regression model relies on a solid foundation of quantitative empirical data. Such models have been criticized for relying too much on data, and not enough on theory, a situation which can lead to models which are useful for predicting in a strict mathematical sense without allowing any theoretical insights to be reached (Howell, 2002). However, when combined with a judicious quantity of theoretical grounding, regression models can be powerful predictive tools, particularly in applied settings.

Schubert, Friedmann & Regenbrecht's path-analysis model (1999)

Schubert, Friedmann & Regenbrecht present a path-analysis model of presence which is a good exemplar of its class. The path analysis technique is a method related to statistical regression which is used to find and test complex causal relationships amongst networks of observed (measured) and latent (unobserved) variables (Loehlin, 1998). Schubert *et al*'s model is extraordinary in its class, as it has an extremely strong conceptual basis. At the same time, its path model was derived from a large sample, giving it a significant capacity for making quantitative predictions.

Schubert *et al* base their model partly on a cognitive theory outlined by Glenberg (1997 in Schubert, Friedmann & Regelbrecht, 1999), which rejects the notion of direct perception. Schubert uses this theory to argue that interaction with an environment would be impossible if only the information present in the optic flow field were available (the basic tenet of the direct perception school). Instead, supplementary information must come from a representation of the environment held by the user (a similar concept to Craik's *mental model*). This representation contains possible patterns of behaviour which are suitable to that environment. These actions are prioritized, based on survival needs and other relevant criteria, so that action in the environment always occurs in a meaningful and goal-directed manner. Schubert *et al* argue that this basic representation structure applies not only to operation in real environments, but also to virtual environments. However, virtual environments differ from real environments in that the operator in a real environment is not confronted with the problem of two information sources, as all information comes from the real environment. Users of virtual environments thus have to engage in two processes; the first is constructive (processing information from the virtual environment), and the other suppressive (suppressing contradictory signals from the real environment, such as tripping over cables, the low resolution of displays, etc). From these two processes, argue Schubert *et al*, arise a set of possible behaviours which make sense for the user to engage in.

Based on these premises, Schubert *et al* factor-analyzed items drawn from seven presence measures, with the aim of determining if this conceptual basis fit into the contemporary body of empirical findings. They conducted two factor analyses, using large samples; 246 participants in the first study, and 296 in the second (details of these studies can be found in Schubert, Friedmann & Regelbrecht, 2001). From these analyses, Schubert *et al* identified eight independent factors which compose these measures. These factors were (Schubert, Friedmann & Regenbracht, 2001):

1. *Sense of being in a place (SP):* This is the commonly referred to aspect of presence ("being there"), as referred to by Steuer and Slater & Usoh.

2. *Involvement/attention (INV):* This includes items related to attention and concentration. Schubert *et al* take the appearance of this factor as evidence of the importance of suppression of real world stimuli to presence.

3. *Realness (REAL):* This factor related to subjects' opinions of the reality of the VE. Schubert *et al* argue that this factor occurs as a consequence of judgments made by the subjects, rather than due to features of the VE itself. They argue that it is probably not a part of the presence experience itself, but only a co-varying event.

4. *Quality of immersion (QI):* This refers to the richness and vividness of the display. It is similar to Steuer's notion of vividness.

5. *Perception of dramatic structures and content (DRA):* This factor includes the user's perception of a plot or other dramatic structure in the VE. It also refers to the user's experience of dramatic moments such as unexpected events, excitement, etc. while immersed in the VE.

6. *Interface awareness (IA):* This refers to the degree to which the interface interferes with the user's experience. It also expresses the degree to which the users are familiar with the interface.

7. *Exploration (EXPL):* The opportunity for users to freely explore the virtual environment.

8. *Predictability (PRED):* The user's ability to predict would happen next in the VE (similar to Steuer's motion of mapping).

Having identified these basic factors, Schubert *et al* used the path-analysis technique to build a structure and associated path weights. To create the model, they begin with two basic assumptions, which are drawn from the literature. The first assumption is that the immersion factors (QI, DRA, IA, EXPL and PRED) will be causes of SP, INV and REAL (this is the classic "immersion affects presence" hypothesis which IJsselsteijn *et al*, Steuer and Slater & Usoh express in their models). The second assumption is that SP and INV will also affect REAL; that is, that highly present and involved subjects will regard the environment as being more realistic.

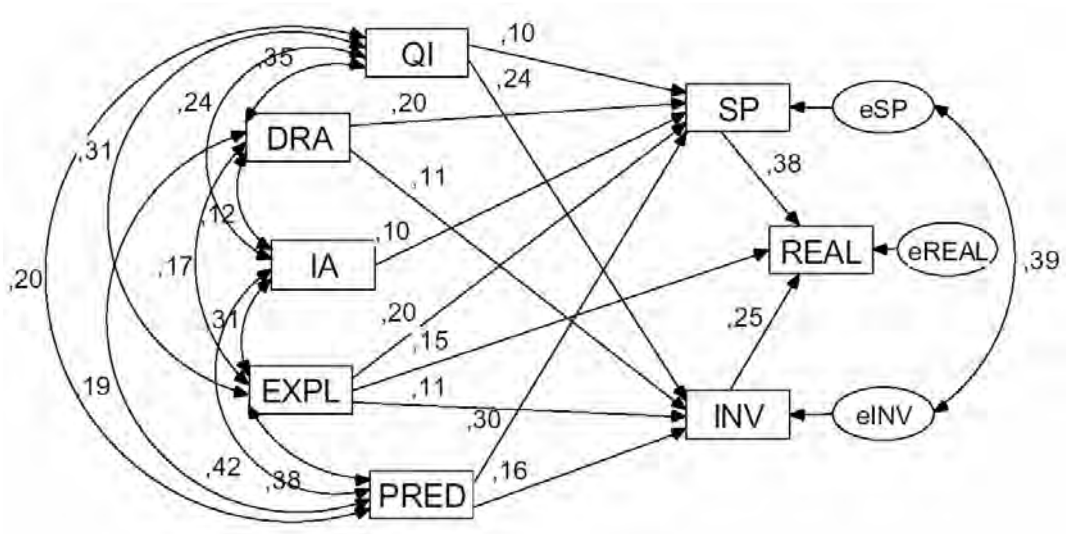


Figure 2-5: Schubert, Friedmann & Regenbrecht's Path Model (from Schubert, Friedmann & Regenbrecht, 1999). Path weights are standardized.

Based on these two assumptions, Schubert *et al* construct the path model illustrated in Figure 2-5. This model shows only the paths between factors which proved to be statistically significant. The model shows several interesting relationships. Firstly, all the predictors affect spatial presence (SP). However, only exploration (EXPL) affects the perception of reality (REAL). Also, both spatial presence and involvement affect the sense of reality (REAL).

Schubert *et al* point to several aspects of the final model which validate their initial conceptual stance. Firstly, exploration and predictability are the two most significant contributors towards spatial presence (with regression weights of 0.2 and 0.3, respectively). This supports the notion that spatial presence arises not only from the information contained in the optic flow field, but also from the user's experiences in the environment. Secondly, the perception of dramatic content also contributes significantly towards spatial presence (with a weight of 0.2); this is taken as evidence that the meaningfulness of a space contributes towards the degree of presence that is felt in it.

The Schubert, Friedmann & Regelbrecht model is quite impressive for several reasons. Firstly, it has the potential to be used as a tool for the quantitative prediction of presence; this would be of inestimable value for both researchers and engineers. Secondly, the model exists with a sound theoretical basis, which has been empirically supported, and is thus suitable for expansion by further examination of the underlying theory. Finally, the model includes many aspects from previous presence models. It thus presents the opportunity for previous presence related findings to be reinterpreted in its terms.

Unfortunately, the model has faults as well. Schubert *et al* criticize their own methodology on the basis of using the same set of data for the factor and path analyses. This error violates the fourth of Hughes' (1994) guidelines for predictive models, which states that a model should be constructed and verified with independent datasets. A second problem of this study is the lack of validity analysis on the scales used to produce it. If a model is to be successfully created mostly from data, it is extremely important that it possess high levels of construct validity (Anastasi & Urbina, 1996). This is because statistical models find relationships between *scores*, not between *variables*. If a test does not have construct validity, then a score obtained on that test does not reflect the true level of that variable. A model based on unvalidated scales simply predicts test scores rather than true phenomena. Indeed one of the assumptions required by the path-analysis technique is that there be no measurement error (Loehlin, 1998), which is highly unlikely in this case. Because this problem lies with the currently available measures of presence, this criticism does not only apply to Schubert *et al*'s model, but to all statistical models which make use of current scales.

Chapter 3

An introduction to schemata, scripts and connectionist architectures

This chapter introduces three of the central psychological ideas used in the construction of the connectionist model of presence, namely schemata, scripts and connectionist architectures. The understanding of these concepts has been developed by cognitive psychologists since the mid 1950s, and so there is a great deal of empirical evidence to support their existence and structure. This chapter briefly explains each concept, and reports some of the more seminal or compelling studies which provide empirical support for their existence. How these concepts can be applied to the problem of explaining presence will be descried in chapter 4.

3.1 Schemata and scripts

A person's knowledge is often thought of as simply a collection of *object* concepts (such as *computer*, *chair*, *coffee*, etc.). While it is true that any individual's knowledge contains thousands of these object concepts (Eysenk & Keane, 1995), an important part of knowledge exists in the form of *relational* concepts, which describe the relationships between object concepts (for instance, *type, sit* and *drink*). More recent theories of human knowledge, such as those of Collins & Loftus (1975) have concentrated on understanding the relationship between object concepts, relational objects and how they are encoded in the mind, by modeling them as networked architectures. In these models, called *semantic networks*, object concepts are expressed as nodes, with relational concepts expressed as connections between nodes. Thus, two connected nodes have a relationship between them (such as the relationship *sit* exists between *writer* and *chair*), and a lack of connection between two nodes means the two objects are not related (such as *computer* and *coffee* might not have a relationship connecting them directly).

Research into these semantic networks uncovered some evidence to support them. Gentner (1981), for instance, gave her subjects three types of sentences to remember. The first type (which she called the *general type*), were vague, and thus each node included connections to many other nodes. In the second type (the *poorly connected specific type*) each node contained connections to only a few nodes, but of those, some were unrelated to the object of the sentence. The third sentence type (the *well connected specific type*) contained only a few connections between nodes, and most of them were directly related to the object of the sentence. Gentner found that the poorly connected specific type were the worst remembered type of sentence (because a large proportion of the existing connections led to irrelevant objects). The best remembered were the well-connected specific sentences (where a large proportion of the existing nodes were connected to relevant objects). The general sentences performed slightly better than the poorly-connected specific sentences. This may come as a surprise, as these types of sentences contained less information about the event depicted in the sentence than the poorly connected type. The effect occurs because as there are so many connections between the verb and other objects, a high proportion of these were interpreted as being relevant to the object of the sentence. This study emphasized the idea that relational knowledge could not be modeled by a simple function of the amount of information available, but rather required more complex models such as semantic networks.

Further research into the semantic network model revealed that the mental representation of relationships was more complex than was previously realized. Coleman and Kay (1981) showed that verbs can have different types of relationships, and also create implications about the properties of the subject and object they can take. They give the example of the verb *to lie*, which they argue, implies three distinct properties of its subject and object. Firstly, it implies that the statement made is false. Secondly, it implies that the subject believes the statement is false, and thirdly, it implies that the subject intended to deceive. This complexity cannot be easily expressed in a semantic network.

This type of evidence combined, in the mid 1970s, with a realization by researchers that explaining knowledge about events and other complex concepts required more than information about the attributes of objects. Even positing a hierarchy of relationships, as Gentner did, was not enough to solve the problem. Impelled by the work being done in linguistics by theorists such as Chomsky (1965) on recursive models in language rule encoding, researchers such as Rumelhart (1975) and Thorndyke (1977) proposed "story grammars" which were though to underlie the encoding and comprehension of stories, as stories usually which contain a complex set of relationships between objects. The pinnacle of this line of inquiry was the creation of two closely related theoretical constructs – the *schema* and the *script.*

3.1.1 Schemata (Rumelhart & Ortony, 1977)

A schema is a recursive, organizing structure which encodes complex concepts by combining simple objects with the relationships that occur between them (Eysenck & Keane, 1995). Schemata contain a list of *relations* and *variables* on which these relations act. The relations can encode simple transitive verbs (such as *sit* or *drink*), or they can encode complex relations indicating causation and temporal relationships (such as *cause* or *produce*). The variables can contain either concepts (such as objects) or other schemata; however, before an object or schema can be assigned to a variable, it must satisfy some test condition. These tests prevent nonsense constructions such as *eat the air*. These variables contain default values, so that if enough information is not provided to assign a value to all the variables, a complete, workable schema is still available. For example, if a subject is given the sentence "Joe drank heavily", which does not contain information about what Joe drank, the information is still available by virtue of the default value. In this example, most subjects would respond to the question "what did Joe drink" with *alcohol, liquor* and other related terms, as these pass the test to fit into the *drinks heavily* schema. It is unlikely that objects such as *water* or *milk* would fit that schema.

Schemata encode generic knowledge, which can be applied to a range of situations, provided those situations do not violate the variable fitting test conditions. For example, the relation CARRY can apply to *Joe carried the bottle* as well as it does to *The ship carries the passenger*, but not to *The passenger carries the ship*, as this violates the necessary test condition (which might be *SUBJECT must be larger than OBJECT*). The ability of schemata to encode generic situations allows for greater efficiency in memory, as few schemata can be used to encode a wide variety of situations. Also, the existence of default values of variables allows for optimal processing to continue even under conditions of reduced information. However, the information provided by the default values might be incorrect, as it is based upon previous experiences in similar situations (Rumelhart & Norman, 1981). This can lead to errors in behaviour or cognition, as the information used for decision making does not fit the environmental situation the person is facing.

3.1.2 Scripts (Schank & Abelson, 1977)

Schemata, as defined above, are capable only of encoding statements. They are clearly not suited to encode more complex, sequence dependent behaviours which are common in everyday living (such as *posting a letter* or *having coffee*). A hoary example of such a complex action is having a meal at a restaurant. Operating in this situation requires following a strict sequence of events: waiting to be seated, selecting from the menu, waiting for the meal, eating, waiting for the bill, paying and leaving. In order to successfully carry out a trip to a restaurant, it is necessary not only to have knowledge of each step, but also of what order they should be performed in. It is also necessary to be able to identify the signs that one step is completed and that the next has commenced. Clearly, a schema alone cannot encode all of this information. Schank and Abelson propose a *script*, which is a structure that encodes a stereotypical sequence of events. Scripts are said to exist for sequences of events which are experienced frequently; for instance, most pedestrians have a *crossing the road at a traffic light* script, but scripts can also encode the far more obscure – long time fans of *Star Trek* probably have a script for getting a starship underway.

As the examples above show, scripts exist to co-ordinate knowledge and behaviour in order to achieve complex goals such as eating in a restaurant while not causing a scene. In order for a script to encode all this information, it is necessary for the script to encode not only the actions required, but also the relationship between behaviours (such as temporal order) and the required role-players (such as the waiter). Each script is represented by a series of variables referred to as *role-slots* (Schank & Abelson, 1977). Role slots can be filled either by a role (such as waiter) or by a schema (such as entering the restaurant, or ordering the food). Based on the specific situation at hand, role-slots will be filled by the objects and information present in the environment. For instance, the *waiter* role-slot might be filled by the teenager behind the counter wearing the restaurant's logo on his t-shirt, or it could also, in a different situation, be filled by a tall, elegantly dressed middle-aged man holding a pen and notebook. Not any object can fill a role-slot, however. Each role-slot has criteria which are applied to determine if a particular object is suitable for filling the role-slot (Schank & Abelson, 1977). These criteria are stored in long-term memory, and are acquired through personal experience and other forms of learning, such as observation. If a particular person has only been to five star restaurants, it is likely that for this person a youth behind a counter would not satisfy the criteria for *waiter*. These criteria ensure behaviours are kept appropriate to the environment, while ensuring storage efficiency by allowing adaptation of scripts to new but similar situations.

It seems that the purpose of scripts is to allow for algorithms of cognition and behaviour to be applied to achieve particular goals in an environment. These algorithms are applied only to situations which fit a set of criteria for suitability, through a mechanism which involves the use of perception (to asses the resources currently available in the environment) as well as previous knowledge (to provide criteria to compare against, as well as to determine if the available resources are similar enough to those required). By means of this mechanism, behaviour in environments can be kept appropriate to the situation, which in turn ensures that the required goals will be met. The inclusion of learning to modify the scripts allows for behaviours to remain goal directed (as the basic structure of the script remains unchanged), but still relevant to environments which change (by adapting the criteria for accepting an object or schema as a role-slot filler).

A satisfactory amount of evidence has been gathered to support the theory of scripts. The commonality of script content was demonstrated by Bower, Black & Turner (1979). They collected descriptions of restaurant eating experiences from a large, diverse group of subjects. Although there was a great deal of variation in terms of the participants, the types of restaurant they frequented and the restaurant experiences they had had, Bower *et al* discovered a common core of 15 ordered experiences which were common. While this study demonstrates the content of scripts, and the phenomenon of time order, Galambos & Rips (1982) demonstrated the effect of scripts on general mental processing. Firstly, they primed a script in their subjects by showing their subjects various photographs of restaurants. Then, they asked subjects to decide if particular actions were part of the experience of visiting a restaurant.

Galambos and Rips found that subjects took significantly longer to make the decisions when the action was not part of the script, than when it was. This demonstrated the effect of scripts not only on regulating behaviours, but also on making particular types of information more available.

3.2 Connectionist architectures

Psychologists began to use theoretical models of computing machines as an aid to explaining cognitive processing in the mid 1950s (Martindale, 1990). This movement (referred to as the *information processing paradigm*) applied theoretical computing machines (such as the Von Neuman machine) to attempt to replicate some of the effects seen in human cognition. The method achieved some notable successes, such as the attention model of Triesman (1964) and the measuring of short term memory capacity (Miller, 1956). However, some of the fundamental principles of the information processing paradigm were found to be lacking by some researchers. For example, many theories based on information processing required the existence of a *central executive*, a hypothetical processing unit which makes basic decisions. Critics of this approach, such as Paivio (1975), claim that postulating the necessity for such a device makes the very information processing approach moot as an explanatory tool. The claim is that rather than explaining a phenomenon, the information processing approach merely re-phrases it. The problem in question is simply transformed from a situation of having to infer what is happening in the mind, to one of having to infer what is happening in the central executive. The mystery and obscurity which the explanation was meant to dispel, is simply passed onto the central executive. An true explanation of a phenomenon, maintain these critics, must leave no “black boxes” which defy explanation, unaccounted for.

3.2.1 Basic properties of connectionist architectures

A connectionist architecture defines a computing machine without recourse to any “black boxes”. The components of the machine are simple and easily understood, and its computational capacity arises mostly from its structure rather than from any componential complexity. Various authors have written about connectionist architectures, and have called them various things, ranging from “unorganized machines” (Turing, 1948 in Teuscher, 2001) to “parallel distributed processes” (Rumelhart, Hinton & McClelland, 1986). Although the terminology changes, a few basic characteristics of connectionist architectures are largely agreed upon:

1. *A group of nodes* (also referred to as “processing units”). The only functions a node can perform are to take on an activation level, and to release that activation at a particular time to the other nodes to which it is connected (Martindale, 1990).

2. *Activation state*. Each node can take on a different level of activation. Activation of a node beyond a certain threshold leads to awareness of the concepts encoded by that node. Most authors allow only limited amount of activation to be spread throughout the network, thus limiting the maximum processing capacity of the network (Rumelhart *et al*, 1986).

3. *A set of connections between nodes.* Each node is connected to at least one other node by a connection which can either be excitatory or inhibitory. Activation received through excitatory connections increase the activation level of a node, while inhibitory connections decrease the activation level. Each connection represents a long-term memory for an association between two concepts (Martindale, 1990).

4. *Unit functions for the nodes.* Each node contains two functions which it applies to activation. The first (the *input function*), is used to scale the inputs received from particular connections. The input function also determines the rate at which activation in the node decays. The second

function (the *output function*) determines the activation level at which the node will commence transmitting activation to the nodes it is connected to (Rumelhart *et al*, 1986). The input and output functions of a node are always different; if they were not, the node would simply acts as a pass-through gate (Martindale, 1990).

5. *Learning rule.* In a connectionist architecture, learning simply means modifying *connection strength* (i.e. the ease with which the connection transmits activation). Several rules have been proposed, but the most commonly used is that of Hebb (1949), which states that a connection's strength will increase if both connected nodes are simultaneously activated.

6. *System Environment.* Most cognitive researchers provide for a particular structure to the system. Fodor (1983) argues that nodes are clustered into *analyzers* each of which is specialized to perform a specific sub-task in the system. These analyzers are massively interconnected to allow communication between themselves. Inside each analyzer, further structure exists. Theorists such as Grossberg (1980) and Martindale (1981) propose that nodes inside analyzers are arranged into various layers. The connections between nodes inside each layer are generally inhibitory, while those between nodes of different layers are generally excitatory.

3.2.2 The interactive activation and competition architecture

From the basic architecture described in 3.2.1 above, a large number of system variations have been developed for various purposes. Caudill & Butler (1990) describe many of these variations such as crossbar networks, Kohonen networks and Local Linear models. Each of type of network is suited to a particular class of processing task (Caudill & Butler, 1990), and as such not all will be of relevance to this work. We will focus the discussion on one particular type of network, the *interactive activation and competition architecture*, which is the most used model in psychological modeling (McClelland & Rumelhart 1986). This model has been successfully applied and developed by many researchers (Martindale, 1990). The rest of the discussion on connectionism in this work will refer to this type of network.

The interactive activation and competition architecture forms the basis of a *cognitive analyzer*. An analyzer (sometimes called a *module*) is a unit of processing which is takes as input a set of data (from other modules), and processes it to output more abstract data (Martindale, 1981). Many analyzers of human cognition have already been identified (Koronoski, 1967). An example of such an analyzer is the unit which recognizes faces. This analyzer takes as input simple features (lines, curves, regions of shade, etc) from other perceptual analyzers and outputs a single percept (a face). The identification of modules is generally done through a process of inference based on clinical data. For instance, it is known that face perception exists in a isolated module separate from other forms of perception, as damage to particular regions of the brain leads only to prosapagnosia (the inability to perceive faces), while brain damage leading to other forms of perception deficit will not affect face perception (Eysenk & Keane, 1995).

3.2.3 Structure of the analyzers

The microstructure of the various analyzers is different, but all share the same basic macrostructure (Martindale, 1981). This structure is illustrated in Figure 3-1. Four basic structural features exist in each analyzer: Stratification, vertical stimulation, lateral inhibition and similarity mapping.

Stratification

Each analyzer contains a large number of nodes, arranged into four or more layers (Martindale, 1990). Each layer represents a particular level of abstraction. Lower layers represent more basic concepts (with the bottom-most layer receiving the information to be processed from other analyzers), with upper layers representing more abstract concepts (with the top layer outputting the processed information to the other analyzers).

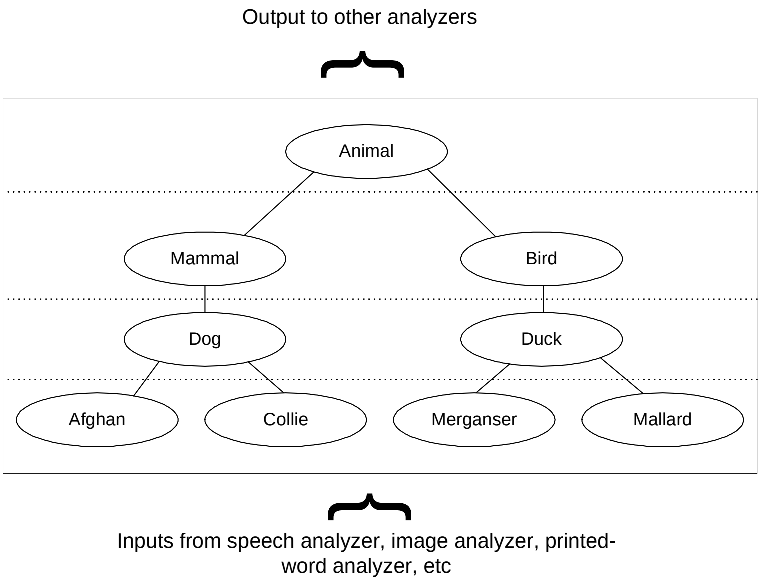


Figure 3-1: An example of analyzer structure. This is a semantic analyzer, with only a few nodes shown in each layer for clarity (from Martindale, 1990)

Vertical Stimulation

Connections exist between nodes on a particular layer and nodes on adjacent layers. These vertical connections are always excitatory, and serve to encode mostly *is a part of* relationships (Martindale, 1990). For example, in Figure 3-1, **Dog** *is a part of* **Mammal** *is a part of* **Animal**. These connections are bidirectional, insofar as they are able to transmit activation upwards and downwards.

Lateral inhibition

Bidirectional connections also exist between nodes belonging to the same layer. Unlike the vertical connections, these lateral connections are of the inhibitory type (Martindale, 1990). The existence of these connections is a point of some contention, although they are widely recognized as being necessary in these models. Theorists such as Grossberg (1980) argue that inhibitory connections of this type are essential, as a network with only excitatory connections, such as those proposed by Beurle (1956), will tend to either end with all nodes maximally activated, or, once the activation has decayed, with zero activation.

Similarity mapping

A final property of these networks, proposed by Kohonen (1989), is similarity mapping. On each layer of the analyzer, the nodes are arranged so that nodes encoding similar concepts are situated close to each other. This arrangement allows for nodes encoding similar concepts to inhibit each other strongly. This prevents ambiguity, as the inhibition makes it unlikely that two nodes encoding similar concepts would become highly activated simultaneously.

3.2.4 The spread of activation and resonance

Because the connections in interactive activation and competition architectures are bidirectional, there is not a strong distinction between input and output. It is quite common for feedback to occur, and many of the unique features (such as learning) can only occur if activation spreads backwards from a top-level node to a lower node - a process known as backpropagation (McClelland & Rumelhart, 1988). To more fully understand these feedback loops, it is necessary to distinguish between two classes of activation spread, namely top-down activation, and bottom-up activation (Eysenk & Keane, 1995).

In the case of bottom-up activation (also called data-driven processing), the network receives activation primarily from sensory or less processed sources (Martindale, 1990). However, often the senses do not always provide sufficient information to allow complete processing of the percept. Often, bottom-up activation leads to ambiguity as many nodes in a particular layer become activated, and due to their inhibition of each other, the situation that results is a continuous oscillation of activation between nodes, which provides no clear outcome to the process (Martindale, 1990). To resolve the deadlock and allow processing to continue, top-down activation is required. Top-down activation (also referred to as conceptually-driven processing) originates as abstract thoughts or expectations, and is transmitted downwards by means of the bidirectional vertical connections. This extra activation can increase a node's activation state to the point where it is able to overcome the inhibition by its neighbours, and break the deadlock.

The situation described above, where top-down and bottom-up activation converge, leads to a state known as *resonance*. Resonance occurs when top-down and bottom-up activation meet at a particular node. In sensory terms, this occurs when what is being observed matches that which the observer was expecting to see. When such a match occurs, the activation continues to be transferred along the path joining the source of the activation, leading to further inhibition of neighbouring nodes. Resonance acts as a mechanism of signal amplification, by inhibiting "noise" nodes and increasing the activation in the resonance or "signal" path.

3.2.5 Evidence for the importance of top-down activation in perception

Top-down activation plays an important role in all cognitive processing, not least of all in perception. During the early 1960s, the *direct perception* school (whose efforts are exemplified by the work of J.J. Gibson) argued that perception can be accomplished only with the information that is available from the senses, and that no top-down activation was necessary for successful perception (Eysenk & Keane, 1995). This idea received a great deal of research attention, and, as will be seen below, the evidence collected suggests that regardless of whether the senses provide enough information or not, top-down activation plays a significant role in perception.

The first example of the importance of top-down processing can be found in the *word-superiority effect*, which is the phenomenon where letters are more easily processed if they are embedded in words than if they stand alone. Reicher (1969) demonstrated this effect in a simple experiment. Reicher showed subjects either real or fabricated four letter words for an instant (for instance *walk* and *wlka*), and then presented them with a single letter. The subject's task was to decide if the letter existed in the word they had just been shown. Reicher discovered that subjects viewing the real words made the decision far faster than those viewing fabricated words. This can be easily explained by top-down processing, as knowledge of the word would allow the subjects to more easily identify which letters it was composed of.

A similar effect was found by Pollack and Pickett (1964), which has been dubbed the *sentence superiority effect*. Pollack and Pickett recorded natural conversations between people, and then cut single words out from the recordings. They played these single words to subjects and asked them to identify the words. Only 47% of the words were identified by the group of subjects. This implies that even if all of a piece of speech is heard clearly, many of the words are simply unintelligible and thus carry no meaning. However, sentences are usually understood completely (if not, clarification is normally requested by the listener). These results show that some of the information for processing sentences must come from a source other than the senses.

To demonstrate that top-down processing is not only useful in speech processing, Williams and Weisstein (1978) conducted an experiment which led to the coining of the general term *object superiority effect*. Williams & Weisstein briefly showed their subjects images of a line with a reference dot. The lines were chosen from only one of four possibilities, and the subject's task was to identify which of the four possibilities they had seen. Williams & Weisstein found that subjects identified the lines far more accurately when the stimulus lines were shown as part of a meaningful image compared to when the stimulus lines were presented in isolation. They concluded that the context of the image made the processing and subsequent memorization of the line far easier. This demonstration proved to be most damaging to the direct perception claims, as the visual senses carry the most information of all the senses; yet Williams & Weisstein managed to show that if top-down information was available, it enhanced processing, even of visual stimuli.

Chapter 4

Cognitive presence

The discussion of presence conceptualization in Section 2.3 in Chapter two shows that the presence field does not yet have a single, agreed upon definition of presence; indeed, in their recent review of the presence literature, Schuemie *et al* (2001) identify more than ten major conceptualizations of presence in use. Lombard & Ditton (1997) attempted to reduce this complexity by creating six categories of presence definition. In Section 2.3 of Chapter two, we attempted a similar feat by grouping definitions into three broad classes. Regardless of how one attempts to cluster definitions together, this does not overcome the problem of the variety of measurement methods which exist. This chapter considers the various conceptualizations of presence and presence measurement philosophies outlined in Chapter two, and attempts to derive a unified concept of presence by making use of psychological research into environmental processing (discussed in Chapter three). We also provide a measurement method for our conceptualization. Finally, we consider some of the strengths and weaknesses of our approach.

4.1 The need to unify presence

There are several reasons why it would be desirable to have a single conceptualization of presence. Lombard & Ditton (2000) identify two. The first is the lack of standard measuring instrument for presence. Several other researchers, such as Slater (1999), have expressed dissatisfaction not only at the quality of contemporary presence measuring tools, but also at the number available (see section 2.4 in Chapter two for a review of some of these measures). The second reason identified by Lombard & Ditton (2000), is that there exists no standard measurement methodology; some researchers, such as Witmer & Singer (1998) use a post-experience, introspection methodology, while others, such as Freeman *et al* (2000), perform naturalistic observations of the subjects during the VE experience. Although it is true that some of this variety arises simply from a desire to explore methodological alternatives, a great deal of it seems to come from differences in the basic ideas of what constitutes presence.

To these two compelling reasons for unifying the concept of presence, we add a third. A single common conception of presence would allow for an integration of presence research findings. This would permit researchers to meaningfully compare findings from various studies, which would in turn allow for the creation of theories with wide scope. For example, Schubert, Friedmann & Regenbrecht (1999) attempted to create a statistical model to predict presence by combining various presence questionnaires. Although they added into their model a commendable variety of presence concepts, they were not able to include them all. This is because the basic method used by Schubert *et al* (1999) requires quantified measurements; however, some presence approaches, such as those which observe reflexes or spontaneity of behaviour, are not amenable to quantification, and it was thus impossible to include them.

4.1.1 Lombard & Ditton's unified concept of presence

Some work has already been done towards creating a single, unified concept of presence. The best known attempt at unifying the various conceptualizations of presence was made by Lombard & Ditton (1997). They identified six major movements in presence conceptualization (which they refer to as *dimensions of presence*). These were:

1. Presence as social richness (the degree of interaction with other users in the VE, ease of communication, etc)
2. Presence as realism (the fidelity of visual, auditory and other stimuli)
3. Presence as transportation (the sense of "being in" another place)
4. Presence as immersion (surrounding the user with stimuli from the VE while blocking stimuli from the real world)
5. Presence as social actor within medium (the user's sense that they belong socially in the world by means of storyline etc)
6. Presence as medium as social actor (the user's sense that the VE is able to sustain social intercourse)

Lombard & Ditton argue that although these factores seem quite disparate, a common factor can be extracted from them. They propose that the six categories all include the idea of *the illusion of non-mediation* to some degree (see 2.3.2 in Chapter 2 for a further discussion of this concept). They argue from that if one were to create for users the illusion that the VE medium were invisible, then each of the six properties would consequently increase, and conversely, that as each of the six dimensions increases, so too would the strength of the illusion.

A notable problem with Lombard & Ditton's argument is that it fails to establish convincingly that *the illusion of non-mediation* is the only common thread between all six presence dimensions. Lombard & Ditton do present a convincing reason as to why non-mediation is related to each of the six dimensions they define. However, in none of the six cases do they establish that non-mediation is the sole cause, or that non-mediation is itself not caused by a more general construct.

The objection we present above is one which could in principle be applied to any definition of presence which aims to become a bridge between previous concepts. There is a danger that instead of simplifying the concept of presence, such a move would make it more complex. If it should transpire that the *illusion of non-mediation* is not the atomic mental operation which produces presence, then Lombard and Ditton's ideas would have effectively transformed presence (a single term) into a hierarchical structure, a result which is in direct opposition to the central aim of replacing many concepts with a single one. There is not enough evidence at the moment to determine if Lombard and Ditton have created the most inclusive concept of presence; indeed it would be impossible to show that a definition of a non-trivial concept is all-inclusive. Our objection is not that Lombard and Ditton have not created the most inclusive concept possible, as this objection could only be supported by a great deal of empirical evidence, which we do not have. Rather, our objection is that they have created a conceptualization which *allows the possibility* of having a more abstract cause, and thus introducing more complexity.

4.1.2 The psychology of behaviour regulation as a unifying principle

Based on the arguments presented in 4.1.1 above, we suggest that presence cannot yet be regarded as a successfully unified concept. We therefore propose our own concept to unify presence, while remaining mindful to avoid the objection we raised to Lombard & Ditton. To do this, we will consider presence not as a special unique experience, but rather as a result of normal psychological processes.

Slater & Usoh's concept of presence (which is "the (suspension of dis-) belief that [the subjects] are in a world other than where their real bodies are located", 1993, p.222) suggests that presence is a novel experience which is only possible when a person is viewing a virtual environment. Similarly, Zeltzer (1992), who suggests that presence is the state which arises when we are immersed in a high bandwidth stream of information together with interaction, provides a view of presence which we feel suggests presence is a unique experience which is only possible due to the existence of virtual environments. Indeed, even the more technical views of Sheridan (1992, 1996) contain this view. Sheridan seems to consider presence as a phenomenon which "pops up" when the correct causal circumstances have been arranged. We feel that considering presence as a unique psychological phenomenon in this way is incorrect, because the writings of Slater & Usoh, Zeltzer and Sheridan do not contain an analysis of how they arrived at the conclusion that presence is a unique experience; rather, it seems that they engage in conceptual reification; that is, they assume that presence is a unique phenomenon, without first considering the possibility that presence may simply be a special case or unfamiliar form of an already well-understood mental process.

Our own conception of presence is far more similar to those views presented by Schloerb (1995), Thie and van Wijk (1998), Lombard & Ditton (1997) and especially Zahoric & Jenison (1998). Like these theorists, we believe that the first step towards understanding presence lies in discovering if presence does not already exist in some other body of literature. Thie & van Wijk, for instance, make reference to the human factors literature by defining presence in terms of active mental models. Similarly, Zahoric & Jenison refer to environmental perception theory to explain presence. For each of these groups, presence is simply the application of existing mental processes to a virtual environment. This pragmatism has several advantages. Firstly, if presence is defined as new concept, then one must begin studying it from the ground up. However, if presence is defined as a special case of an existing field, then it is only necessary to determine how presence differs from the other phenomena in that field. A second reason, suggested by Schloerb's *subjective presence* notion (1995), is that working with a new concept leads to imprecise measurements of it. Because a new concept has no theory behind it, it becomes difficult to determine the construct validity of its measures (Anastasi & Urbina, 1996). Finally, considering presence as a new phenomenon can lead to theoretical fragmentation and unnecessary complexity. As Lombard & Ditton (1997) show, considering the various strands of presence research as special cases of a larger structure can lead to a deeper understanding of it.

We propose that presence can be understood as the consequence of psychologically processing any environment, including virtual environments. This point is in line with the views of Slater & Usoh (1993), Schloerb (1995), Lombard & Ditton (1997), and Sheridan (1999). We therefore propose that presence can be best understood by considering the underlying cognitive processing of the virtual environment. As demonstrated in Chapter three, cognitive psychology contains a thorough understanding of the relationship between environment, cognition and behaviour, and this understanding is based on empirical evidence collected over a period of more than forty years. We feel that this field can contribute a great deal, both in terms of theory and findings, towards the understanding of human behaviour and cognition in virtual environments.

We are certainly not the first in the presence field to suggest that cognitive psychology can be applied to understanding user behaviour in virtual environment. Bystrom, Barfield & Hendrix (1999), for instance, postulate that the direction of attention can have an impact on presence. Also, Mania & Chalmers (2001) use memory and recall performance to determine presence. Finally, Freeman, Avons, Meddis, Pearson, & IJsselsteijn (2000) propose a presence measurement strategy which expects users to apply familiar, real world behaviours and cognition strategies when they are present in the virtual environment. Judging from the successes of these researchers, we feel confident that cognitive psychology can be fruitfully applied to understanding presence.

Our concept of presence relies on the notion of schemata (Rumelhart & Ortony, 1977) and scripts (Schank & Abelson, 1977). As stated in Chapter three, scripts and schemata are theoretical structures which encode complex behaviours which are initiated in response to particular conditions, some of which are perceived, and some of which are remembered. The perceived conditions are similar to those used by the behavioural presence movement (e.g. Loomis, 1992). However, unlike the idea of "reflex actions" used by Usoh, Arthur, Whitton, Bastos, Steed, Slater, & Brooks (1999), which are triggered only by perceived stimuli, scripts and schemata require a complex set of conditions existing in the mind of the user. This is more akin to the spontaneous social behaviour which Sheridan (1996) suggests as a sign of presence. Our contention is that presence is the state when the scripts and schemata active in the user are those which apply to the virtual rather than the real environment.

In this light, we can re-examine the findings of Usoh *et al* (1999), who report that when presented with a virtual hole, users react by refusing to step over it, even though they knew that no real consequences could arise from this act. Usoh *et al* (1999) argue that when presented with the virtual hole, subjects can respond in one of two ways. The first (non-present) way is to simply "disbelieve" the hole and step through it; the second (present) way is to treat the hole as real and refuse to step through it. We propose that each of these response behaviours (which are far more complex than simple reflexes) is in fact encoded by a script. The first response (which Usoh *et al* would call "non-present") is encoded by a script which is appropriate in the real world, as there is no hole in the floor of the experimental venue. The second response (the "present" response) is encoded by a script which is appropriate to the virtual environment, where the hole actually does exist. Whether the subject is present or not, in our view, depends not on what they believe is true about the world (a volitional action), but rather on what script their mind executes (an automatic action).

The script concept can also be applied to explain why subjects show social behaviours, such as greeting, to computer controlled agents, as reported by Cassell, Torres, & Prevost (1999). Again, there are at least two scripts which can be applied to the situation: the script appropriate to the real environment (not to greet the agent, as greeting only applies to real persons), or the script appropriate to the virtual environment (to greet the person in front of you). If presence simply means the selection script appropriate to the virtual environment, then we can expect that present subjects will greet the agents, while non-present subjects would not.

We thus propose the use of script and schemata theory as a unifying concept for presence. We feel that using this field overcomes our objection of Lombard & Ditton's *illusion of non-mediaiton*, because scripts and schemata are well researched, and are considered to be atomic structures of thought perception (Eysenk & Keane, 1995). Scripts and schemata have been used to explain a variety of complex behaviours, including social behaviours (such as restaurant attendance - Bower, Black & Turner, 1979), and should therefore apply to most situations which could occur in a VE. Also, our use of a well-researched field minimizes the risk of reification. Finally, because this field has a long research tradition, there exists an established research methodology, which in turn allows for more robust research to be conducted (Eysenk & Keane, 1995).

4.2 Cognitive presence

Based on the arguments presented in 4.1.2, we propose a novel conception of presence, which we call *cognitive presence*, which we believe can be used as a unified concept of presence. Cognitive presence is defined as *the degree to which the virtual environment dominates over the real environment as the basis for thought* (Nunez & Blake, 2001). This same basic notion has been recently echoed in Slater (2002b), in which the author states, "[t]he critical issue is how will the actor respond? To which set of signals will the actor respond?" (pp. 436).

In our definition, we consider only the abstract part of the virtual environment for this purpose, and not the technology used to display it. By this we mean that we consider not the VE as it is rendered by the display, but rather the VE as it is stored in the VR system. This excludes any display artifacts which may occur during display as not being part of the virtual environment. We believe that display artifacts are aspects of the real environment, and not of the virtual environment, and thus should not be considered to be contributing towards cognitive presence. We define a virtual environment as *data displayed in such a way as to create in users the impression of objects existing in a space* (Nunez & Blake, 2001).

4.2.1 Immersiveness and cognitive presence

In line with current research (which we describe in Chapter two), we consider the technology used to display an environment to be a determinant of how a subject mentally processes that environment (for some examples of this evidence, see Stark, 1996, Slater, Usoh, & Chrysanthou, 1995, and Barfield, Rosenberg, Han, & Furness, 1993). However, we decide not to explicitly include any notion of display technology or immersion in this definition. We agree that there is a strong theoretical basis supporting the importance of display technology for presence. For instance, the theories of Steuer (1992), Slater & Usoh (1993), Schubert, Friedmann & Regenbrecht (1999) and of IJsselsteijn, de Ridder, Freeman & Avons (2000) emphasize the importance on display technology. However, the theories of Thie & van Wijk (1998), and the Immersion Presence & Performance model (Bystrom, Barfield & Hendrix, 1999), while recognizing the importance of display factors, emphasize other aspects of presence. Indeed, even the theories of Steuer and Slater & Usoh mentioned above accept that display factors are only a part of the workings of presence. This lack of consensus on the importance of display factors is not the sole reason for our decision to exclude them from our definition, however. We feel that the emphasis on display variables in presence research is not due only to their actual importance, but also to the weight of research which has been directed at immersive displays. The lack of evidence on the importance of non-immersive technologies does not indicate that these technologies cannot produce presence, but rather that this issue has not been sufficiently researched (see Towell & Towell, 1997, for a notable exception). Due to this bias in the evidence, we decide to not include display variables explicitly in the definition, as this would preclude the investigation of cognitive presence in non-immersive displays.

4.3 The Contents of Consciousness Inventory (COCI) - a measurement of cognitive presence

To conduct quantitative research with cognitive presence as a variable, a measure is required. We created the Contents of Consciousness Inventory (COCI) as a measure of this construct. In Sections 4.3.1 and 4.3.2 below we discuss the problem of measuring cognitive presence, and describe the measure. In Section 11.3 we present a psychometric evaluation of the COCI, and consider its effectiveness as a measure of presence.

4.3.1 The problem of measuring cognitive presence

Cognitive presence is defined in terms of the rules or basis which the mind uses to process information obtained from the senses. It is therefore not possible to infer the degree of cognitive presence simply by considering the environmental stimuli. What is required is a method of establishing which script the user is applying to process the environment. The solution to this problem lies in the fact that scripts affect not only the selection of behaviour, but all cognition in general. For instance, in Galambos & Rips's study (1982), the activation of the *attending a restaurant* script led to an effect on an abstract word processing task, which is far removed from the behaviour required to successfully attend a restaurant.

The generality of this effect can be exploited for the purpose of measurement. We propose a method where the subject is presented with a list of stimuli, and their task is to select one stimulus from the list.

If one of the stimuli in the list is related to the script which is active in the subject, then that stimulus should seem to stand out from the others. In Galambos & Rips's study, for instance, words which were related to restaurants were identified much faster than those not related to restaurants. We thus infer that in a list of items, those items which are related to the script which is currently active in the subject will seem different from the others. If the item selected from the list is related to the script which would be appropriate in the VE, then, according to the definition presented in 4.2 above, the subject is present.

A complication in applying scripts and schemata to measure cognitive presence lies in the fact that when a user is involved with a virtual environment, there are two possible sources of environmental information. Witmer & Singer (1998), Bystrom, Barfield & Hendrix (1999), IJsselsteijn, de Ridder, Freeman & Avons (2000) and others have theorized that presence can only occur when the user's attention is focused on the VE display. However, Triesman (1964) and others have shown that attention does not shift discretely, so that it is possible to experience *divided attention*. In this case, the environmental processing mechanisms of the mind will be receiving input from two sources simultaneously. However, the script literature suggests that only one script can be active at a time (Schank & Abelson, 1977). We suggest that under such conditions, the user might be sometimes more focused on the VE and at other times more focused on the real environment. To capture a single presence score under these conditions, we implement a strategy similar to that used in the *Breaks in Presence* method of Slater & Steed (2000). They solve a similar problem by asking subjects to report when they experience a *break in presence* (i.e., a time when they become aware that they are no longer focused on the VE, but rather on the real environment) at any time during the VE experience. At the end of the VE session, the subjects which have experienced fewer breaks in presence are deemed to have experienced more presence. To overcome the divided attention problem, we propose sampling the mental state of the subjects at various points over time during the VE session. This overcomes the potential problem of the subject not focusing attention at the precise moment of measurement, as other opportunities for measurement will exist during the time of the experience.

This method also overcomes two difficulties experienced by the currently popular introspection methods discussed in section 2.4.1 of Chapter two. Firstly, measurement occurs during the VE session, and thus overcomes any memory distortions and other related difficulties. Secondly, the method does not require subjects to introspect in any way; their task is simply to select an item from a list. This prevents any of the negative introspection effects suggested by Nisbett & Wilson (1977).

4.3.2 A first measure of cognitive presence - The Contents of Consciousness Inventory (COCI)

We implemented the measurement philosophy discussed in 4.3.1 above. The measure we created, which we call the Contents of Consciousness Inventory (COCI), uses word lists as stimuli. The COCI is computerized, being administered by the machine displaying the VE during the VE session, with one item appearing approximately once every minute. The exact amount of time between COCI items is randomized slightly (by as much as 10 seconds) to prevent creating in the user an expectation of the appearance of the next item. When the item is displayed, the user is briefly shown (for less than half a second) a word fragment. The subjects are then shown a list of four words, and their task is to select from the list the word which they believe the word fragment represented.

The word fragment task is designed to be completely ambiguous, as any of the words from the list could fit the fragment. The task is in fact a ruse, designed simply to create a pretense for the word selection. The word fragment task aims simply to provide a motivation for the subjects to process the words in the list so that the script-induced cognitive bias can take effect.

The COCI score is derived by summing the number of times the user selects a word in the list which is related to the script which would be active if the VE were determining the activation of scripts. We designed the word lists so that in each item only one word would be suitable for selection. We did this by creating VEs with a particular theme (a hospital and a medieval monastery), and adding in each item one word which was related to the VE setting (for instance, for the hospital items we added the words *ambulance* and *antiseptic*). Scoring the COCI is therefore trivial; because the themed words are already known, it is a simply case of scoring 1 for the item if that word was selected, or 0 if it was not. Technical details of how the COCI was implemented in each of our VR systems are in sections A.6 and B.5 of Appendices A and B respectively. The list of COCI items we used are in Appendix F. A discussion of the effectiveness of the COCI as a presence measure may be found in sections 11.3 and 11.5.4 of Chapter eleven.

4.4 Critique of the cognitive presence approach

The cognitive presence approach conceptualizes presence as the effects of the virtual environment on basic thought processes. This level of abstraction permits one to try to explain current presence findings from a unified perspective. This is possible because we consider what we believe is a common cause. For instance, presence findings based on perception (e.g. Hendrix & Barfield, 1995 or Lessiter, Freeman, & Davidoff, 2001) are separate from those findings based on posture and movement (e.g. Freeman *et al*, 2000). This separation exists because there is no common theoretical ground joining the two approaches. However, if one considers the common basic cognitive processes underlying each of these phenomena, it might be possible to find a framework or theory which incorporates both of these classes of finding, and explains how they relate. This type of integration is the aim of the cognitive presence approach.

To achieve integration in this way, the cognitive presence approach needs to be fairly general. This generality may lead to a degree of inclusiveness which some researchers might argue extends too far. For instance, under the definitions presented in this paper, a novel or other piece of text could produce cognitive presence. If the text defines a space with objects included in it, the reader might, using their imagination, become cognitively present in that space. This inclusiveness comes about mostly because, like Lombard & Ditton (1997), we do not explicitly define any display medium criteria for achieving presence. This position opposes the idea, expressed by Slater & Wilbur (1995), that presence requires immersion as a necessary condition. We do not consider this position a weakness of our definition. We do not believe that there is enough firm evidence to exclude non-immersive displays as potential inducers of presence. Indeed, Johns, Nunez, Daya, Sellars, Casaneuva & Blake (2000) showed that presence can be successfully elicited using non-immersive desktop diplays. More impressively, Towell & Towell (1997) showed that presence can be elicited using text-only displays. We therefore leave the relationship between immersiveness and presence as a matter of investigation, rather than one of definition.

A second, more serious weakness of the cognitive presence approach is that it does not explicitly include the concept of user-environment interaction. Interaction is seen by many researchers as essential for presence. For instance, Zeltzer (1992) requires that some form of interaction with the environment occurs before a user becomes present. Similarly, Waller, Hunt & Knapp's (1998) notion of *simulation fidelity*, requires that interaction takes place before a VE can be said to be having an effect on the user. However, the term "interaction" itself is not as straight-forward as one might think. Sheridan (1992) as well as Witmer & Singer (1998), for instance, consider simply moving in an environment to be a form of interaction with it. On the other hand, Zeltzer (1992) considers interaction to be a changing of the state of the VE. To further confuse matters, Regenbrecht & Schubert (2002) present empirical evidence of how the user's belief that interaction with the VE is possible can be enough to increase the degree of presence. Due to this theoretical murkiness, we decided not to include interaction as a component of cognitive presence, again opting to leave that as an issue for further empirical investigation.

Chapter 5

The connectionist model of presence

Chapter 2 outlined some of the current major trends in presence modeling. This chapter presents our own model. The model which we propose in this chapter is based on McClelland and Rumelhart's (1986) interactive activation and competition architecture, and includes aspects of Schank & Abelson's (1977) script theory. The model implements an environment processing architecture, the purpose of which is to use both perceptions of the environment as well as previous experience and abstract knowledge to select and maintain cognitive strategies and behaviours which are appropriate for operation in that environment. This model thus aims to explain presence in terms of the cognitive processing of virtual environments, and the resulting cognitions and behaviours therein.

This model is most similar to Schubert, Friedmann & Regenbrecht's (1999) model, although it differs in many respects. While Schubert *et al* created the structure of their model via factor and path analyses, we modify an existing architecture (interactive activation and competition architecture), based on current presence research. This includes many of the causes included in Ijsselstein, de Ridder, Freeman & Avons' model (2000), as well as in Steuer's model (1992). Like Slater &Usoh (1993) we use a strong foundation in a previous field, but in our case it that foundation is found in experimental cognitive psychology.

We propose a model with three layers of analyzers. The topmost of these layers (which we call the *conceptual layer*) represents abstract, higher-level conceptual processes, and the bottom layer (which we call the *perceptual layer*) represents data-driven perceptual processes. The middle layer, (which we call the *action layer*) represents an intermediate stage of environmental perception, where previous knowledge and expectations converge with current perceptions of the environment. It is in this middle layer that we propose the key to understanding presence lies. Each of these layers, as well as their interconnections, is explained below. The basic structural layout of the model is presented in Figure 5-1 on the following page.

The reader should note that the connectionist model currently exists as a qualitative tool only; it is not capable of predicting presence numerically (as is possible with Schubert & Regenbrecht's model). Rather, the model exists to relate the effects of perceptual stimulation with those of the user's mental state in order to show how these two forces combine to generate a sensation of presence in the user. It is possible to show how a particular set of circumstances would produce a strong or weak sense of presence, but not to predict the degree of presence in a quantitative sense. Section 12.6.5 and 12.6.6 in Chapter 12 contain a discussion of how this basic model can be expanded to provide quantitative predictions of presence.

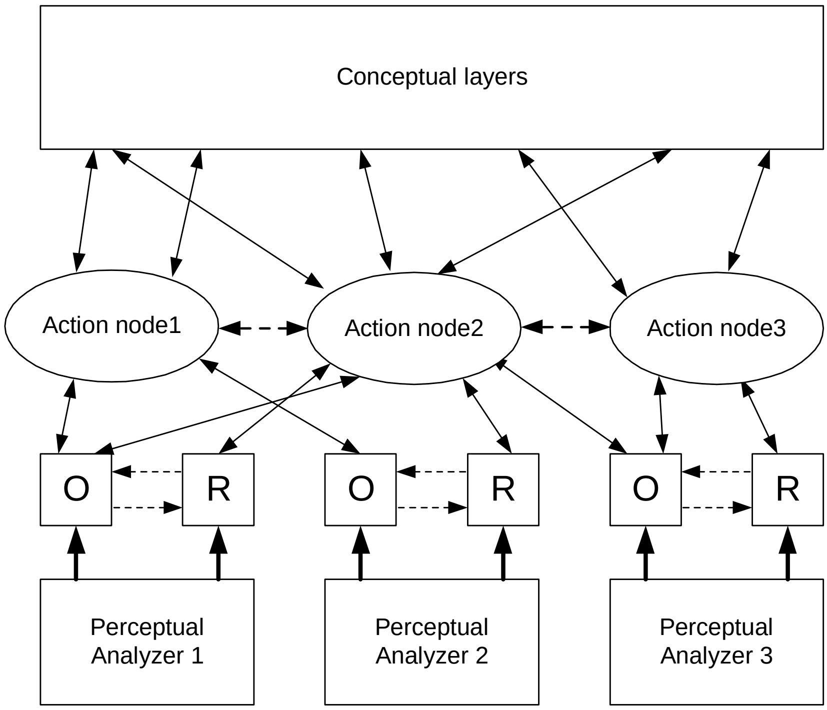


Figure 5-1: The basic structure of our connectionist network. The model consists of three horizontal layers: at the bottom, the perceptual layer (which consists of the perceptual analyzers and the O and R nodes), in the center the action layer (which consists of the action nodes), and at the top the conceptual layers (which are regarded as a single functional layer by the model). This figure only shows three analyzers per level for the sake clarity, but the model includes many more in each level.

5.1 The perceptual layer

The perceptual layer plays similar role to The external component of Slater & Usoh's model (1993), and the bottommost layer of Steuer's model (1992). We propose that this layer represents a series of perceptual analyzers, there being at least one per sensory modality. The need for separate analyzers per modality is derived in part from the ideas of Slater (1999), who suggests that presence is a function of many separate display system factors. Slater argues that many of these factors are independent of each other, and this independence allows VR systems engineers, who generally work with limited display resources, to trade factors off against each other to maximize presence in a particular system. In order to model this independence, it is necessary to have separate perceptual analyzers, so that the level of activation in any one analyzer can be held constant while the activation the others can be changed.

For example, Slater's argument implies that a VE which includes both sound and graphics might produce different presence levels to one which only includes graphics, and some evidence to this effect is presented by Hendrix & Barfield (1995b).

By modeling the perceptual analyzers of sound and vision as separate entities, it is possible to explain this phenomenon by arguing that in the graphics-only VE, only the vision analyzer is activated, while in the graphics and sound VE, the graphics analyzer remains activated, but the auditory analyzer becomes active as well. It is this increase in overall activation which leads to the increase in presence.

The top level of each perceptual analyzer contains two nodes (labeled 'O' and 'R' in Figure 5-1) – one of these nodes (the 'O node') will become activated if the perception by that modality results in the percept being identified as an *object*, and the other (the 'R node') will becomes activated if the perception by that modality results in the percept as being identified as a *rendition of an object*. We define a rendition of an object as any symbol or reproduction which aims to represent or symbolize an object, regardless of how physically dissimilar the rendition is from the object. In the case of vision, renditions include photographs, video footage, geometric models, sketches and even a verbal description of a scene can all be regarded as renditions of a scene. The modeling of the O and R nodes is derived partly from the work of Lombard & Ditton (1997) who argue that presence is the illusion that the VE experience is not mediated by technology. In this model, mediation would be experienced when a the R (rendition) node is highly activated, leading to the perception of the objects not as real, but as representations of real objects. On the other hand, if the O node is highly active, this will lead to the perception of the virtual environment of consisting of real objects, and consequently, the sense of non-mediation.

The O and R nodes are connected to the action layer by means of excitatory connections. Each O and R node is connected to many action nodes, with each connection encoding one of the possible responses which are appropriate for any particular percept. Activation of an O or R node will lead to a partial activation of all the nodes in the action layer to which it is connected. Usually, the activation provided by the perceptual layer is not sufficient to activate the action nodes to a level where a response occurs. In order for that to occur, a certain amount of activation is also required by from the conceptual layers. For example, a user who perceives an object as a door in a virtual environment may respond to that object in one of many possible ways, none of which is directly suggested by the percept itself, but rather by previous experiences as well as by abstract knowledge about doors (Norman, 1988). In a few cases, such as reflex or automatic actions, perceptual stimulation alone will provide sufficient activation of the action layer for a response to occur.

5.1.1 Further motivation for modeling the O and R nodes

Slater & Steed (2000) emphasize the difference between experiencing data originating from the "real world" (which they term being in an R state), and experiencing data emanating from a VE display (which they term being in a V state). Specifically, they note that data from a VE display will contain glitches in image quality, sudden changes in frame rates etc. This type of display artifact clearly creates a distinction between VE data and "Real world" data, but we think the distinction is more subtle that "real world"/VE. Our decision to create a separate apparatus for detecting objects and renditions of objects is based largely upon the work of Gibson (1979). Gibson suggests that the rendition of an object is not simply a poor quality reproduction of an object, which contains less information than the object itself. Rather, Gibson points out, a rendition is often identified as such by the fact that the rendition contains information which the object itself does not. For example, we can appreciate the difference between an apple and the photograph of an apple not simply due to the failings of the photograph to capture all the information present in the apple (such as depth information, its hardness, smell, etc), but also due to the fact that a photograph of an apple will also contain features *of the photograph* such as glossy reflections, the grain of the film, and perhaps the smell of photographic chemicals, which the object does not contain. Gibson further points out that it is possible to fool the visual system into perceiving a rendition when it is in fact an object on display. As an example, Gibson presents a study in which researchers prepared a window of a house so that a majority of subjects reported it as being a framed photograph of a garden rather than a garden itself. Based on these results, we reject the notion that a continuum of quality or amount of information exists between an object and a rendition of that object, rather suggesting that each of these is perceived separately. We propose that the process of deciding if a percept is an object or a rendition occurs as one of the final steps of perception (due to its relatively high level of abstractness), and we thus place it at the top of the perceptual analyzer. Furthermore, we allow the 'O nodes' and 'R nodes' to connect to different action layer nodes, based on the notion that the response to an object will tend to be different to the response to the rendition of that object. For instance, on perceiving an apple (object), my first reaction might be to smell it, but on perceiving a photograph of an apple (rendition of the object), my first reaction would probably not be to smell it, but rather a different response.

This same idea is recently expressed by Slater (2002b) as the notion that presence occurs as a process of selecting between two competing hypotheses: that the percept is either representing a real scene, or a virtual scene. In Slater's view, the scene presents evidence to the viewer which may or may not convince them that the scene they perceive is real. Our conceptual layer can be thought of as an explanation of the mechanism by which this happens. The O node represents the hypothesis that the scene (or a particular aspect of the scene) is real, and the R node represents the hypothesis that the scene is not real. The level of activation in each node represents the degree of confidence that the hypothesis is true, and the final determination of which of the two is true occurs as a process of the activation and mutual inhibition between the two nodes.

As Slater (2002b) reminds us, it is usually not the case that a percept is identified simultaneously as both an object and a rendition, as this would lead to ambiguity and confusion. For example, a wax apple (rendition) might be mistaken for an apple (object), but once it has been identified as a wax apple, it is not likely be perceived as a real apple. This disambiguation feature of perception is important, and we model this by placing the O and R nodes on the same layer, and allowing inhibitory connections between them. These connections allow disambiguation by allowing the node which is more activated to inhibit and thus overpower the other, ensuring that at any moment, one of the two is dominant over the other.

5.1.2 Implications of the perceptual layer model

a) Using an interactive activation and competition model allows for perception to contribute to presence by degrees rather than in a dichotic "on/off" style. Evidence that perception affects presence in this way is contributed by Lessiter, Freeman and Davidoff (2001), who conducted a study into the effect of varying the quality of audio playback on presence. They created four levels of audio quality, and found that presence increased in each subsequent level of quality. This suggests that presence varies continuously as a function of the perceptions by a single modality. The variable activation connections between each perceptual analyzer and the action layer in this model allows for the replication of this effect.

b) This model allows modeling of the benefits of multimodality to presence (as demonstrated, for example by Salln?s (1999), who found that presence in a multimodal environment was superior to that in a unimodal environment) by modeling each of the sensory input channels as a separate, independent, perceptual analyzer, each of which is independently capable of contributing activation to the action layer.

c) This model also explains how a stimulus which is initially perceived as a rendition can, with increased exposure, be later perceived as an object. This effect is demonstrated by Nass, Steuer & Tauber (1994). They asked a group of subjects to engage in a text-based dialog with a computer agent, aware that the agent was only a program, and not another human user. After a period of interaction with the agent, Nass *et al*. found that subjects' interaction style changed, behaving towards the agent as if they were interacting with another person (for instance by using polite expressions, etc). This finding suggests that although the subjects began by perceiving the program as simply a rendition of a conversation with another person, exposure to the system changed the perception to that of an actual conversation with a person. This type of effect is modeled by postulating separation between the O node and the R node, each with its own connections to the action layer, and by postulating an inhibitory connection between the O and R nodes to allow disambiguation between of the two.

d) Each of the O and R nodes has its own connections to the action layer. This models the possibility of different reactions based on whether the stimulus is perceived as an object or a rendition. This effect can be seen, for example, in the difference in postural reactions which can be caused by changing from a simple projected display (which is more likely to be perceived as a rendition) to a stereoscopic projected display (which is more likely to be perceived as an object), reported by Freeman *et al*. (2000).

5.2 The conceptual layers

Like many theorists, we believe that the user's psychology has an important role to play in the experience of presence. The conceptual layers represent the more abstract processes of cognition. The conceptual layer has the same purpose as Ijsselstein, de Ridder, Freeman & Avons' (2000) *user characteristics*, as Steuer's (1992) *willing suspension of disbelief,* as Thie & van Wijk's (1998) *willingness to become present*, as Bystrom, Barfield & Hendrix's (1999) *suspension of disbelief* and part of the *attentional resource allocation,* and as Schubert, Freidmann & Regenbrecht's (1999) *perception of dramatic structures and content.*

For the purposes of this model, it is not essential to describe their workings in detail. The conceptual layers are where mental models of interaction and of environments reside, memories of previous environmental interactions, as well as task-related knowledge and heuristics. The conceptual layers are connected to the action nodes below by means of excitatory connections. When any state of activation exists in the conceptual layers, activation spreads downwards to the action layers, partly priming certain nodes there. For example, a user who encounters a door in a virtual environment will begin thinking about the possible operations which can be performed on the door (Norman, 1988). These include, for instance, opening it, or knocking. In this model, the user's previous experiences with doors will lead to a particular higher-level cognitive state, which in turn will contribute activation to a few related nodes on the action layer (for example, the nodes encoding opening, pulling, etc). In most cases the conceptual layer will contribute activation to many action nodes, is essence providing a range of choices which have been appropriate in similar situations in the past. The activation contributed by the conceptual layers is usually not sufficient to activate the action layer sufficiently for a response to occur. For that to occur, some contribution from the perceptual layer is usually required. This requirement for an external stimulus ensures that the selected behaviour matches the environment in which it is going to be expressed.

5.2.1 Motivation for modeling the conceptual layers

The conceptual layers are included in this model of presence due to a small but compelling body of work that illustrates the importance of complex, abstract cognition on presence. An example is the work of Sas & O'Hare (2001), which shows a relationship between navigation (a high level cognitive activity) and presence. Another example of this is the tendency, reported by Slater, Usoh and Steed (1995), for users who identified with their avatars in the virtual environment, to experience more presence. Apart from the evidence that is available, several influential researchers (such as Sheridan) have indirectly suggested that the user's expectations of the virtual environment will have an impact on the level of presence they experience. Witmer and Singer, for instance, suggest that presence will come about if a user's expectations of their interaction with the VE are met (Witmer & Singer, 1998).

We do not model much detail in the conceptual layers, as there is very little research into the effects of higher-level cognitive processes on presence. Fencott (1999), for example, suggests that the content of a VE might affect presence, but provides no convincing empirical evidence to this effect. Similarly, Sheridan (1992) suggests that the content of a virtual environment might affect the degree of presence felt. Some empirical evidence is provided by Schubert, Friedman & Regenbrecht (1999). They constructed a multiple regression model of various variables associated to presence, and find a significant relationship between *Spatial Presence* (the traditional notion of "being there in the VE") and *Drama* (a measure of the dramatic content and structures in the VE). This suggests that abstract cognitive structures (such as knowledge of dramatic structure) are able to affect presence to a degree noticeable in the population.

We model these layers as a single entity, with many unspecified connections to the action layer. For the time being, we suggest that the contribution of users' previous experiences and their expectations cam be modeled by allowing the conceptual layer to activate nodes in the action layer, and the VE's effect on the user's thoughts and decisions (and other higher-level cognitive processes).

5.2.2 Implications of the conceptual layer

a) Having a conceptual layer which is indirectly (via the action layer) connected to the perceptual layers allows for modeling of Gestalt effects and other phenomena (such as illusions) where top-down activation can affect perception (Neiser, 1967).

b) The indirect connection (via the action layer) between perceptual analyzers and the conceptual layer allows for modeling of environmentally appropriate reactions. An example of this phenomenon is the findings of Usoh *et al* (1999), who found that subjects react to a virtual hole in the VE as they would to a real hole in a real floor; that is by skirting around the edge rather than walking across the open space. In this case, the perception of a virtual hole alone cannot explain why users skirted around it, as there are many possible reactions to perceiving a virtual hole. Perception alone can also not explain why one particular response to the virtual hole was preferred by most subjects. The subjects' behaviour can only be satisfactorily explained by considering the combined effects of previous experience and perception.

c) Placing the conceptual layers on the opposite side of the action layer to the perceptual analyzers allows for the conceptual layers to contribute activation to the action layer independently of the perceptual analyzers. This allows the modeling of perceptual disambiguation by means of previous knowledge, which requires simultaneous activation from both the perceptual and conceptual layers. An example of such an occurrence is the perception of idioms or other contextually laden language constructions (Pinker, 1995).

5.3 The action layer

The action layer consists of a single layer of nodes placed between the perceptual analyzers and the conceptual layer. Each of these nodes contains multiple vertical excitatory connections to both the perceptual analyzers and the conceptual layer. The connections between the perceptual layer and the action layer are broader than those connecting to the conceptual layer, allowing the perceptual layers to exert more influence on the action layer than the conceptual layers. This results in perceptions having more influence than thoughts on behaviour selection, ensuring that behaviour remains appropriate to the environmental situation present. Each node of the action layer also has horizontal inhibitory connections to its nearer neighbours in the action layer.

The nodes of this layer encode the possible relationships between a set of stimuli, and a thought pattern or decision-making process. In this sense, the action layer acts as a repository of schema (Schank & Abelson, 1977) with each schema being encoded by either one or several nodes, depending on its complexity. The nodes are arranged so that nodes encoding similar schemata are close to each other. This is necessary to ensure disambiguation of similar schemata by competition, as the activation of one will lead to the inhibition of its nearest (and thus most similar) neighbours. In this way, the perception of a stimulus is guaranteed to activate only one schema, although the schema that is activated will depend not only on what is perceived (which is bottom-up activation from the perceptual analyzers), but also from the subject's state of mind at the time (which is top-down activation from the conceptual layers).

5.3.1 Motivation for modeling the action layer

The action layer exists as an interface between perception and abstract cognitive processes. It acts as a matching mechanism between mental models or expectations about the environment, and the perceptions of the environment. The action layer allows for different reactions to be initiated in response to the same stimuli, based on the subject's current cognitive state. If the perceptual analyzers were directly connected to the conceptual layers, then a specific stimulus would always bring about the same response, as occurs in a reflex-arc reaction (Martindale, 1991), such as the knee-jerk response. This type of predictable, fixed reaction generally only occurs in response to basic perceptions (such as pain) rather than to the complex perceptual processing which is required in perceiving a virtual environment. It is generally understood that a user's reaction to a stimulus will vary according to that user's goals and expectations (Brown, 1965). For instance, the sight of a raging rapid in a river might bring about an avoidant response in a person, but if that person is keen on a whitewater holiday and suitably equipped with a lifejacket, the response might be quite different. For this reason, we model the action layer as an intermediate step between perception and abstract cognitive functioning. We propose that this middle layer is at the heart of understanding presence, as it can be used to indicate the way in which a stimulus which has been perceived will be reacted to. By examining the pattern of activation in the action layer, it is possible to understand the relationship between a stimulus and the user's higher-level cognitive processes, and as such the action layer can act as an indicator of presence, by exposing the interface between the perception of virtual environments and the effect they have on the user's thoughts and actions.

5.3.2 Implications of the action layer

a) The existence of inhibitory connections between action nodes implies that for any set of circumstances, only a single schema will be selected. This means that for any set of circumstances, a subject is unlikely to become confused as to the course of action to take (confusion being the state where more than one set of actions seems appropriate). According to Carver & Scheier (1998), the reduction of conflicts in the selection of goal-directed behaviour is an important feature of human behaviour self-regulation.

b) The action layer allows the modeling of various appropriate reactions (schemata) to any set of stimuli. The selection of the schema will be based on the extent to which perceptual activation matches with the mental state of the user at the time of perception. This notion is supported by Witmer & Singer (1998), who note that presence will most likely vary as a function of the degree to which the environment responds to and fits with the expectations and mental models of the user.

c) The O and R nodes of a perceptual analyzer are connected to many different action nodes, and so many different schemata can be triggered if a stimulus is perceived either as an object or as a rendition. This explains the phenomenon noted by Steuer (1992), that a realistic rendition of an object will lead to more presence than will a non-realistic rendition; simultaneously, it can also explain why users of non-realistically rendered virtual environments, such as text-based virtual environments, are able to experience presence (Towell & Towell, 1997).

d) McClelland and Rumelhart (1986) state that the relative strength of connection in this type of network can change over time and with learning. If this principle is applied to the connections between the O or R nodes and the action layer, it implies that a stimulus may begin by activating a schema appropriate to an object, but as time progresses, the same stimulus may activate the schema appropriate to a rendition (or vice versa). This will occur as a function of the reinforcement or reward a subject receives for activating a particular schema. Although our model allows for this possibility, the lack of any longitudinal research into presence leaves no evidence to support this idea.

5.4 Amount of activation available for processing

Following the suggestions by McClelland and Rumelhart (1986) as well as by Martindale (1991), we propose that there exists a limited amount of activation in the architecture for distributing between nodes. We propose that the conceptual layers share a single, large reserve of activation which can be used to activate the action layer. Also, each of the perceptual analyzers has its own reserve of activation available for spreading to the action layer. However, the sizes of each perceptual analyzer's activation reserves are not equal; each has a different size, with the preferred modality having the largest reserve. By means of this large activation reserve, the preferred modality is able to exert more influence over cognition and processing than do the other modalities.

5.5 The expression of presence in the model

The connectionist model expresses presence in an implicit, distributed way, similar to that used in Slater & Usoh's representation system theory (1993). In our model, presence is expressed as the activation of nodes in the action layer. This is reminiscent of Thie & van Wijk's (1998) notion that presence is the activation of VE mental models rather than RW mental models.

5.5.1 Defining the existence of presence in the model

In the terms of this model, presence occurs when the set of activated nodes in the action layer is providing an appropriate response to the *virtual environment* provided as a stimulus to the subject rather than to the *real environment* in which the subject is physically located. This implies that presence is the state of mind in which the subjects' thought process is aligned towards the virtual environment. This idea is consistent with our definition of Cognitive Presence (Nunez & Blake, 2001).

The use of the word *appropriate* in the above definition is somewhat problematic, as for all but trivial environments it is difficult to determine what an appropriate set of responses might be. However, it is possible to create a list of appropriate responses based on the requirements of the virtual environment in question. For instance, if a virtual environment has been created as part of a theme park, and its purpose is to scare the users (as in a virtual haunted house, perhaps), then appropriate responses to this environment should be fear, surprise, and so on, whereas laughter would not be an appropriate response. Defining the "appropriateness" of responses by means of examining the specification of the virtual environment in this way has the benefit that, to a limited extent, it defines presence in terms of what is useful in the environment. A virtual environment can then be said to be producing maximum presence when it is creating, in users, the desired responses. This definition also facilitates measurement, as the "appropriateness" criterion provides a baseline against which to compare a user's responses.

5.5.2 The action layer and presence

To understand the significance of the action nodes in presence, it is necessary to note that for each VE situation there is a finite set of actions and cognitions that are appropriate. For example, in a VE of a room with a deep pit in the center, the appropriate behaviours include walking around the hole, stepping carefully, and perhaps peering into the pit. Appropriate cognitions would include judging the depth of the pit, and perhaps fear, if the user has a phobia for heights (notice that these behaviours allow, by observing the subjects, to determine if they are virtually “there” in the room – those displaying these behaviours are present, those not displaying them are not present). No single one of these behaviours is a clear sign of presence on its own, as they are appropriate in several settings. For instance, stepping carefully could also be appropriate in a setting where the user wishes to walk quietly. However, the *related cluster* of these behaviours is indicative of being in a room with a pit. In this model, this clustering of behaviours is expressed in the action layer – each node represents a cluster of behaviours (hence the similarity between action nodes and Schank and Abelson’s schemata). Any single behaviour may exist in a wide number of VE situations, but usually only one cluster of them (one action node) is appropriate in a single VE situation.

From the above discussion, it can be derived that a particular action node (or sub-set of action nodes) can be designated as the “presence node” for a particular VE situation (note that the “presence node” will change from situation to situation, depending on what is appropriate at the time). The reason the activated action nodes can be referred to as the “presence node” is not simply a matter of definition, but is rather due to the effect which the activation of these nodes has on the system as a whole. It is important to remember that an action node is not simply a sink for activation to flow into. Once the action nodes have received enough activation, they begin to transmit that activation vertically to both the perceptual analyzers below and the conceptual layers above. As the activation spreads to these areas, it biases the system towards perceiving and thinking in terms related (that is, connected) to the activated action node. The action nodes thus help to set the “state” of the network as a whole. For instance, a particular action node may set the “behaviour for quiet places” state, while another might set the “exploring a dangerous place” state. However, no “state” as such exists, as the activation levels are continuous rather than discreet. The effect is more of a “bias towards”, and is thus better understood by looking at each action node as encoding something similar to a schema.

Like schemata, action nodes do not simply control behaviour in the motor sense. Because they are connected to the perceptual layers, they also affect perception, leading to perceptions which occur in line with the mental state they impose (a phenomenon know as expectation-based processing after Posner & Snyder, 1975). They action nodes are also connected to the conceptual layers and are thus capable of affecting abstract thought. These effects include, for example, the selection of schema-related coping strategies. It is through this mechanism of action node activation that strategies of thought and action are selected to allow coherent and appropriate operation in an environment. Interpreting presence in this framework is thus simple - if the mental state that is activated at the time of experiencing the virtual environment is more appropriate to the virtual environment rather than to the real environment the VE is displayed in, then the subject is experiencing presence.

Chapter 6

Examples of the connectionist model of presence in use

This chapter presents two examples of the model being used to predict presence levels. These examples show how the model described in chapter 5 can predict, given sufficient information about the perceptual and conceptual starting state, the level of presence that a subject will experience. Both examples are taken from published empirical studies. Using published empirical studies for examples has three major advantages; firstly, the method used has been peer-reviewed and their validity is thus not in question. Secondly, the results provided in published papers represent real results rather than speculation on possible outcomes; and thirdly, comparing the results of well-known studies with the presence levels predicted by the model provides a more convincing argument for the model's correctness.

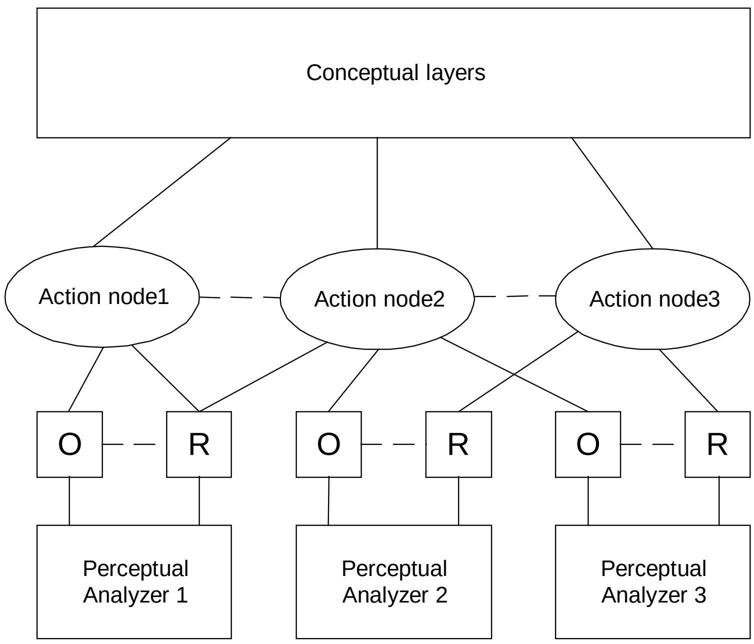


Figure 6-1: The model as used in the examples. In all examples, presence depends on the activation of action node 2.

For these examples, a simplified version of the model presented in chapter 5 will be used. This is presented in Figure 6-1. Nodes which have no activation are drawn with a light line, and activated nodes are drawn with a heavier line. For the sake of simplicity, in each of the examples, we will assume that the degree of presence is largely tied to the activation of action node 2. As described in chapter 5, there is no single action node which can always be said to be the “presence node” – rather, presence will related to different action nodes in each VE situation, and the “presence node” will change from situation to situation.

6.1 Example 1: Modeling the effect of geometric field of view on presence (after Hendrix & Barfield, 1995)

This study investigated the effects of various display parameters, including geometric field of view (GFOV), on presence. Hendrix and Barfield placed subjects into one of three GFOV conditions – 10 degrees GFOV, 50 degrees GFOV and 90 degrees GFOV. The results showed that the subjects in the 50 and 90 degree GFOV conditions reported more presence than those in the 10 degree GFOV condition. This study is typical of a series of studies which show that an improvement in the display capabilities of a VR system leads to an increase in presence (see also Barfield & Hendrix, 1995; Lessiter *et al* 2001 and Usoh *et al*, 1999 for further examples of this type of work). In these studies, generally speaking, no attention was paid to priming the subjects to incite a particular conceptual state, although the perceptual stimulus conditions were carefully controlled.

The connectionist model is well suited to modeling this type of result. We will model two of the three GFOV conditions (10 degree and 90 degree) by varying the relative activation of the O and R nodes of the perceptual analyzer. The 10 degree GFOV condition would be represented by a high level of R activation with a low level of O activation and the 90 degree GFOV condition would be represented by a high level of O activation and a low level of R activation. This is done by assuming that the narrower the geometric field of view is, the more like a generated image the result will look, and the wider the field of view, the more natural and object-like the image will appear. Hendrix and Barfield did not prime their subjects for any particular conceptual state, and so we can assume that subjects entered the experiment with a random set of conceptual node activations. We will model this by assuming that all action nodes are receiving an equal amount of activation from the conceptual layers.

This example will demonstrate how the model can predict the difference in presence comes about between the 10 degree GFOV condition and the 90 degree FOV condition, as reported by Hendrix and Barfield. We will first demonstrate the model predicting a low level of presence in the 10 degree GFOV condition, followed by it predicting a higher level of presence in the 90 degree GFOV condition.

6.1.1 Predicting presence in the 10 degree GFOV condition

Initial perceptual state: We will use perceptual analyzer 2 as the visual analyzer in this example. In this condition, the narrow GFOV would produce a strong frame around the image, which would strongly suggest that the image is generated, rather than natural. This implies that the O node will be receiving very little activation, while the R node would be receiving a lot of activation from the underlying perceptual layers.

Initial conceptual state: Hendrix and Barfield made no attempt to prime their subjects, and so we assume that a random distribution of activation existed in the conceptual layers. We model this by allowing each action node to receive an equal amount of activation from the conceptual layers. These initial conditions are shown in Figure 6-2.

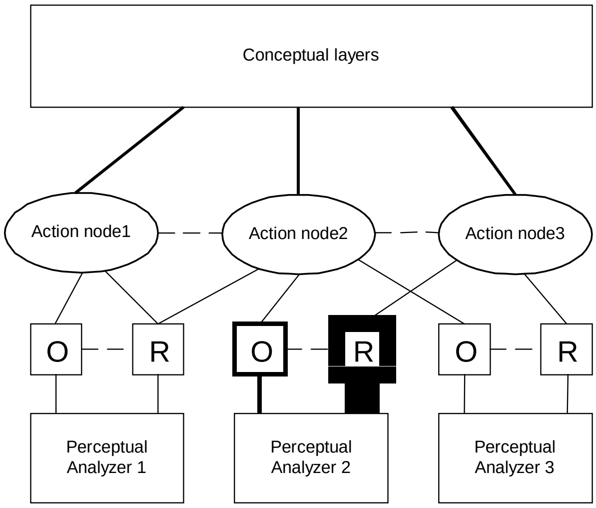


Figure 6-2: Initial state for the Hendrix & Barfield (1995) 10? GFOV example. Perceptual analyzer 2 is the vision analyzer, and action node 2 is the "appropriate action" node encoding presence for this example. Bolder lines indicate more activation.

Once this initial state exists and the model is started, the O and R nodes of perceptual analyzer 2 begin to compete by means of the inhibitory connection between them. The R node is far more activated, and so this dominates and almost completely inhibits the O node, leaving only the R node to make a significant contribution to the action layer. The activation in the perceptual layers will spread along the vertical connections upwards to the action layer. Similarly, the activation from the conceptual layers will spread downwards. This is indicated in Figure 6-3.

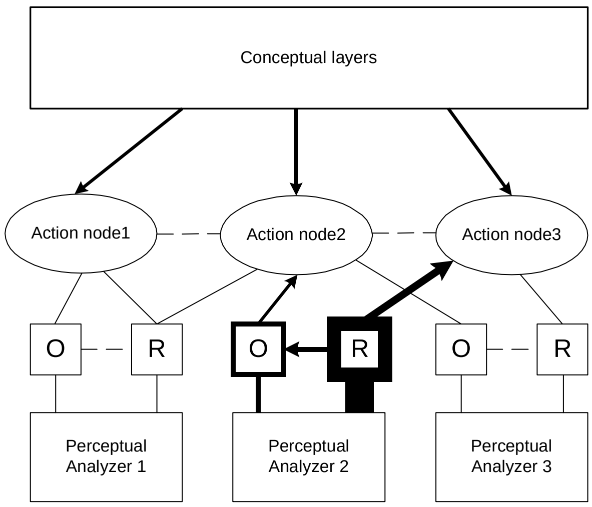


Figure 6-3: Activation spreading in the model

Once the activation reaches the action nodes, they begin to compete and inhibit their neighbours. At this stage, action node 3 is the most activated in the layer as illustrated in Figure 6-4. Although action node 2 is receiving activation from the conceptual layer and a small amount from the O node of perceptual analyzer 2, it is also being inhibited by action node 3, and the net result is a very low activation level in action node 2, with a high level of activation in action node 3.

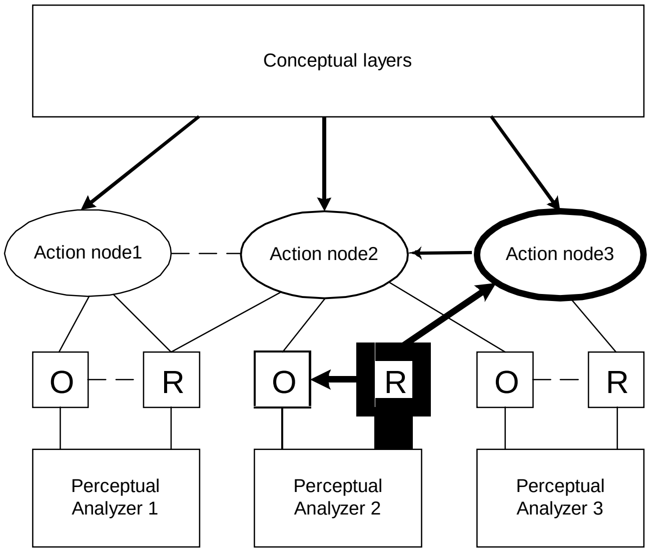


Figure 6-4: Action node 3 receives the most activation in the layer

At this stage, we can see that action node 2 is not very activated, and thus we can deduce that only a small amount of presence will result, by virtue of our definition of action node 2 as the "presence node" for this particular situation.

6.1.2 Predicting presence in the 90 degree GFOV condition

Initial perceptual state: We again use perceptual analyzer 2 as the visual analyzer in this example. In this condition, the wide GFOV would present a more natural image, with less rendition features than the 10 degree GFOV condition. According to the model, the O node would thus be more activated than the R node.

Initial conceptual state: As with the 10 degree GFOV condition, we assume that a random distribution of activation existed in the conceptual layers. We model this by allowing each action node to receive an equal amount of activation from the conceptual layers. The initial conditions for the 90 degree GFOV model is presented in Figure 6-5.

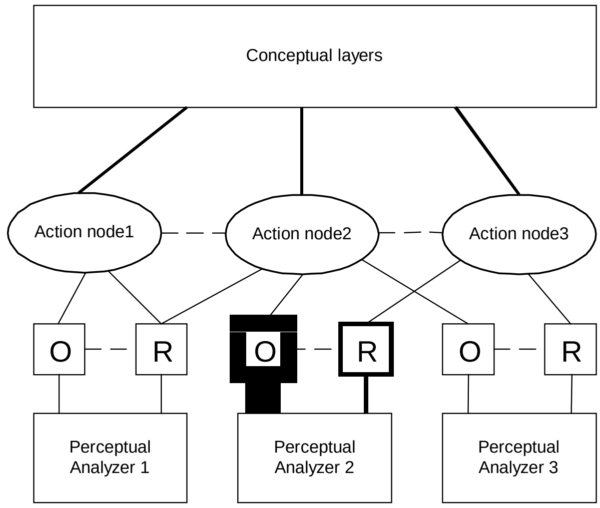


Figure 6-5: Initial state for the Hendrix & Barfield (1995) 90? GFOV example

When the model is started, the O and R nodes of perceptual analyzer 2 begin to compete by means of the inhibitory connection between them. The O node has more activation, and so overwhelms the R node, and leaves itself as the only significant contributor of activation to the action layer. Simultaneously, the activation from the conceptual layers spreads downwards. Action node 2 is receiving a large amount of activation from the perceptual layer below, as well as a slight amount from the conceptual layers above. Because action nodes 1 and 3 are receiving activation only from the conceptual layers, they inhibit action node 2 only a slight amount, but action node 2, which is highly activated, inhibits them greatly. The model will settle with action node 2 containing a substantial amount of activation. This final state is illustrated in Figure 6-6.

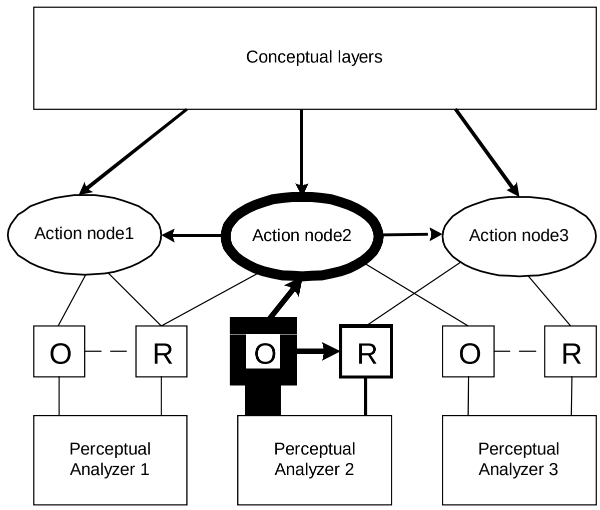


Figure 6-6: Final state of the model

The activation state of the model indicates a high level of activation in the “presence node” (action node 2), and therefore a high state of presence.

6.1.3 Evaluating the model's prediction of Hendrix & Barfield's (1995) finding

The model predicts a higher level of presence (that is, more action node 2 activation) for the 90 degree GFOV condition than for the 10 degree GFOV condition. This prediction matches the empirical finding by Hendrix & Barfield. This shows the model is capable of expressing differences in presence due to a difference in the quality of a single perceptual analyzer. At no stage of the process was there specific mention made of vision by the model, and the reader would be correct in extrapolating that a similar example could be applied to explain the difference in presence levels as a function of audio quality reported by Lessiter, Freeman, & Davidoff (2001). These examples illustrate that this model is able to explain and predict one of the most fundamental and most widely researched aspects of presence, namely that increasing the fidelity of a display variable will increase the amount of presence experienced by a viewer of that display.

6.2 Example 2: Modeling the effect of multi-modality on presence (after Salln?s, 1999)

Salln?s (1999) reports a study comparing the levels of presence experienced by subjects using a VE stimulating a single modality (audio) with those of subjects using a VE that stimulated two modalities (audio and haptic). Two groups of users were asked to perform a collaborative task in one of the two conditions (audio only/audio and haptics), and were measured on various variables, including virtual presence using the Presence Questionnaire (Witmer & Singer, 1998). The results showed a statistically significant difference between the groups, with the group using two modalities experiencing more presence than the one using a single modality. Salln?s concluded from this that multimodal interfaces to VEs lead to more presence than unimodal interfaces.

This result can also be easily modeled using the connectionist model. We will begin by modeling each modality (audio and haptics) as a separate perceptual analyzer. This example will use the same network used in example 1 (presented in Figure 6-1). We will use perceptual analyzer 2 as the audio analyzer, and perceptual analyzer 3 as the haptics analyzer. Again, the "presence node" for this example will be action node 2. As we are not given a measure of quality of the stimuli used by Salln?s, we will assume that the stimuli were of a moderate quality (this is a reasonable assumption, as she found presence levels even using a single modality, which would not have been the case if the stimuli had been of poor quality). We model this by assigning similar amounts of activation in the O and R nodes of each analyzer, with the O node receiving slightly more activation than the R node. Salln?s made no attempt to prime or prepare her subjects cognitively for the study, so we will again assume an even spread of activation from the conceptual layers to each of the action nodes. To model the difference between these two systems we will model each in turn, and compare the amount of activation in action node in each case.

6.2.1 Predicting presence in the audio only condition

Initial perceptual state: We will use perceptual analyzer 2 as the auditory analyzer, and perceptual analyzer 3 as the haptic analyzer for this example. In this condition, the haptic analyzer was not stimulated, and will thus not contribute any activation to the model. As argued in 6.2 above, we will assume similar activation levels in the O and R nodes, with slightly more activation in the O node.

Initial conceptual state: Salln?s made no attempt to prime her subjects, and so we assume that a random distribution of activation existed in the conceptual layers. We model this by allowing each action node to receive an equal amount of activation from the conceptual layers. These initial conditions are shown in Figure 6-7.

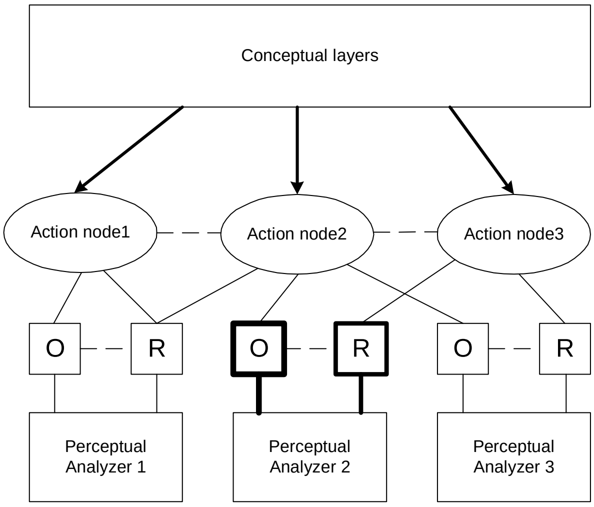


Figure 6-7: Initial state for the Salln?s (1999) audio only condition

Once the model begins to run, the activation in the O and R nodes spreads via the inhibitory link between them, and they begin to compete. As the activation level in each of them is similar, one will not inhibit the other severely. However, the O node began slightly more activated, and thus will tend to slowly overpower the R node.

The O and R nodes will both be contributing activation to the action layer. The O node is slightly more activated than the R node (this difference accentuated by the action of the inhibitory connection between them), and thus action node 2 will receive more activation than action node 3. This is illustrated in Figure 6-8.

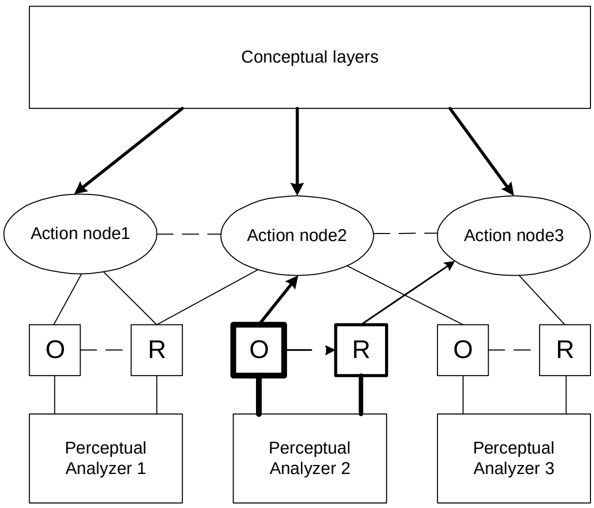


Figure 6-8: Activation spreading in the model

The activation from the perceptual and conceptual layers will begin to spread towards the action layer. Action node 3 is receiving activation from the conceptual layers, as well as a small amount from the R node of perceptual analyzer 2. Action node 2, on the other hand is receiving a moderate amount of activation from the O node of perceptual analyzer 2, as well as a small amount of activation from the conceptual layers. The net effect is a similar activation state in action nodes 2 and 3. However, as the action nodes become activated, they will begin to compete by means of the inhibitory connection between them. As action node 2 is more activated than action node 3, it will exert more inhibition, and will tend to a more active state. The model ends with action nodes 2 and 3 slightly activated, although action node 2 will be more active than action node 3. The final state of the model is presented in Figure 6-9.

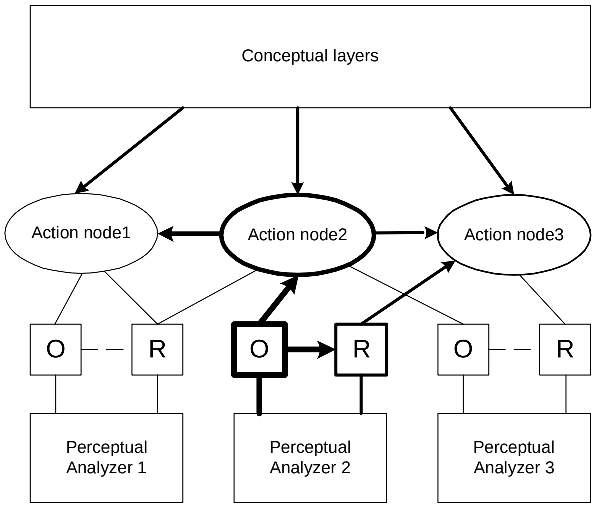


Figure 6-9: Final state of the model

This configuration, as it only has a slight amount of action node 2 activation, represents a moderate presence level.

6.2.2 Predicting presence in the audio with haptics condition

Initial perceptual state: We will use perceptual analyzer 2 as the auditory analyzer, and perceptual analyzer 3 as the haptic analyzer for this example. In this condition, both the haptic and audio analyzers contribute activation, as they are both being stimulated. As argued in 6.2 above, we will assume similar activation levels in each perceptual analyzers' O and R nodes, with slightly more activation in the O node of each.

Initial conceptual state: Salln?s made no attempt to prime her subjects, and so we assume that a random distribution of activation existed in the conceptual layers. We model this by allowing each action node to receive an equal amount of activation from the conceptual layers. These initial conditions are shown in Figure 6-10.

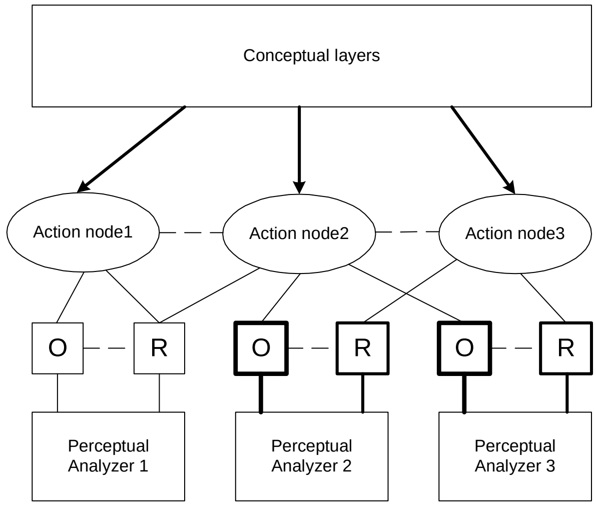


Figure 6-10: Initial state for the Salln?s (1999) audio with haptics condition

As the model begins, the O and R nodes of each analyzer begin to compete via the inhibitory connection between them. As the O node is slightly more activated than the R node (in each case), this will inhibit the other and gain a slight advantage. The activation then spreads upwards into the action nodes. Simultaneously, the conceptual layers contribute an equal amount of activation to each action node. This is illustrated in Figure 6-11.

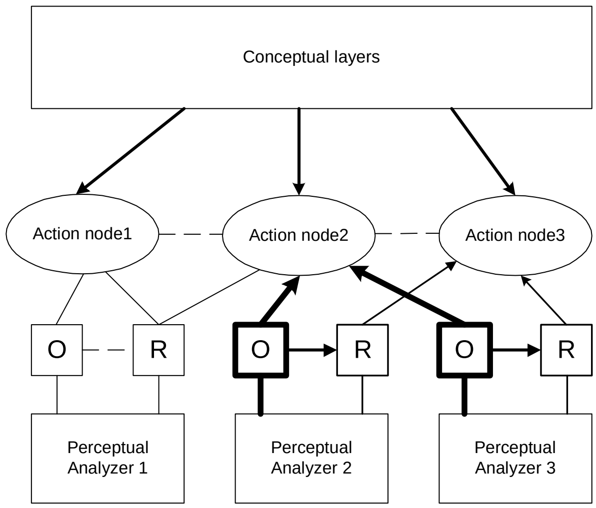


Figure 6-11: Activation spreading in the model

The activation from the perceptual layers contributes activation to action nodes 2 and 3. The O nodes of each will contribute more activation due to their competition with the R nodes, although the R nodes still contribute a small amount. The consequence of this is that action node 2 receives activation from two highly activated O nodes, while action node 3 only receives a small amount of activation from the two slightly activated R nodes. As the action nodes become activated, they begin to compete, and because action node 2 is more activated than action node 3, it inhibits action node 3, and emerges as the most activated node of the action layer. This final state is shown in Figure 6-12.

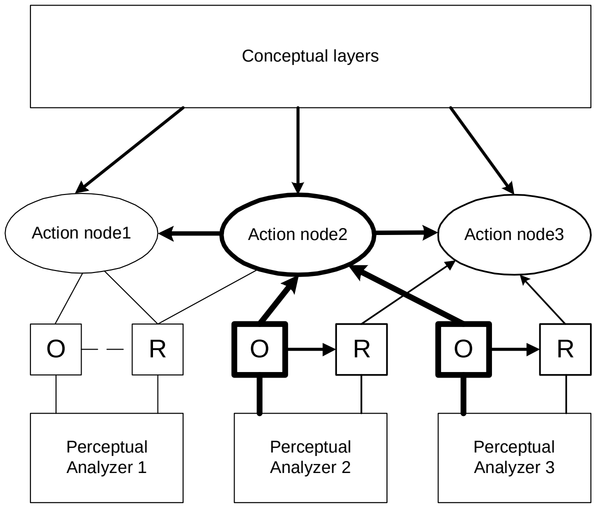


Figure 6-12: Final state of the model

The final state of the model suggest a high level of activation for action node 2 (as it is receiving activation from two moderately activated O nodes), and thus a high level of presence.

6.2.3 Evaluating the model's prediction of Salln?s's (1999) finding

The model predicts a higher level of presence (action node 2 activation level) for the audio with haptics condition than for the audio only condition. This is because in the audio only condition action node 2 was receiving activation only from the O node of perceptual analyzer 2, while in the audio with haptics condition, action node 2 was receiving activation from the O nodes of perceptual analyzer 2 as well as perceptual analyzer 3. This prediction matches the empirical finding by Salln?s. This example also illustrates that the model illustrates the general principle, suggested by Slater (1999), that subjects using multi-modal VR systems will likely experience more presence than subjects using VR systems which stimulate one modality only.

Chapter 7

A strategy for evaluating the connectionist model of presence

Chapter 6 presented how the connectionist model of presence is capable of explaining the results found in two published studies, and arguments were presented as to how these explanations could be extended, by induction, to explain other similar phenomena (such as, for instance, the effects of sound quality manipulation on presence as reported by Lessiter, Freeman, & Davidoff, 2001). However, such post-hoc explanations are not evidence for the efficacy of the model, as these published results were taken into account during the construction of the model, and as such the model agrees with the published findings *a priori*. To show its effectiveness, the model needs to be able to predict a set of results that are beyond the scope of the material used in the model's construction. If the model is capable of predicting the results arising from a novel situation, then the models can be said to be an explanation of the phenomenon it is modeling rather than just an obfuscated summary of previous findings. A perfect model would be able to accurately predict all instances of the phenomenon, past and future, under all conditions.

This chapter presents an evaluation strategy for the connectionist model of presence based on these principles. The model's performance in predicting the published studies presented in Chapter 6 can be taken as evidence that the model is coherent with the current empirical understanding of presence. To convincingly show its effectiveness, it is necessary to create a novel set of initial circumstances, and use the model to predict the outcome. Then, the same circumstances should be replicated in the laboratory, the presence experienced by subjects recorded, and the empirical findings compared with the model's predictions. If the findings match the predictions, then this can be understood as further evidence of the model's correctness.

7.1 Creating novel conditions for evaluation

The model includes two sites for providing input – the conceptual layers, and the perceptual analyzers. By varying the levels of input on each of these input sites, a set of initial activation states can be created. We propose creating four conditions to test, by setting two levels of activation in the conceptual layers, and two levels of activation in the perceptual analyzers.

7.1.1 Manipulation of the conceptual layers

The activation states of the conceptual layers can be manipulated by priming. Priming is a technique used in cognitive psychology to investigate how context affects mental performance, and usually involves exposing the subject to a particular stimulus, not related to the experimental task, which will induce a particular mental state in the subject (Neely, 1977). For example, a subject about to begin a memory task might be shown a series of images before the task, to investigate the effect of the mental context created by those images, on recall.

According to the connectionist model of presence proposed in this document, presence results when the schemata relevant to the virtual environment become maximally active (schemata being expressed in the model by the nodes of the action layer). Any manipulation of the conceptual layers must be made cognizant of the virtual environment in which the subject will be operating. For example, priming a subject with images of bookshelves and books might lead to activation of the schemas associated with operating in a library (including behaviours related to being quiet, browsing, reading, etc). This priming manipulation however, will have different effects on mental performance based on the environment in which the subject is asked to operate. If placed in a VE of a library, for instance, then the priming would contribute towards presence in that environment, as the appropriate “presence nodes” would be receiving top-down activation, and the subject would be more likely to display VE-appropriate (that is, library-appropriate) behaviours. However, if placed into a VE of a football stadium, this manipulation would have no beneficial effect on presence , as the action nodes receiving the top-down activation would not be the appropriate “presence nodes” for the situation.

Based on these considerations, we propose two levels of conceptual layer activation – a *VE relevant* level, in which the subjects are primed with materials directly related to the VE in which they will be operating, and a *VE irrelevant* level, in which the subjects are primed with materials not related to the environment in which they will be operating (it is important to note that, according to Schank & Abelson, some schemata must always be active, so a “no action node active” control is not possible). These two states provide an opportunity to examine the two extremes of the contribution made by the conceptual layers to presence.

7.1.2 Manipulation of the perceptual analyzers

The connectionist model of presence, in its current state of development, describes the perceptual analyzers in far more detail than the conceptual layers, and as such affords many more options for its manipulation. For instance, the number of stimulated analyzers (modalities) could be varied, or the ratio of O to R node activation for a set of analyzers could be varied. We propose to take a broad-brush approach, and create two conditions which represent extremes of activation. In the first condition (*High stimulus quality*) a large number of analyzers are stimulated, with the quality of the stimulus being tailored to produce a high degree of O node activation in each analyzer. The second condition (*Low stimulus quality*), a small number of analyzers are stimulated in such a way as to produce mainly R node activation.

7.1.3 The novel conditions for evaluating the model

The manipulations described in 7.1.1 and 7.1.2 above create four conditions which can be used for evaluating the model. These are *VE relevant priming/High stimulus quality*, *VE irrelevant priming/High stimulus quality*, *VE relevant priming/Low stimulus quality*, and *VE irrelevant priming/Low stimulus quality*. These four conditions permit the investigation of the main effects of each of the manipulations (by keeping the level of one manipulation constant while varying the other), as well as of interactions between the manipulations (by looking at the effect of combining both manipulations at once).

The four manipulations proposed fit into a 2 x 2 factorial design, with conceptual layer initial state as one independent variable and perceptual analyzer initial state as another independent variable and VE appropriate action node activation ("presence node" activation) as the dependent variable. The four cells of the design each describe one of the four novel conditions which will be used to test the model. The factorial design is illustrated in Table 7.1 below.

		Conceptual Layer Manipulation	
		High Stimulus Quality (H)	*Low Stimulus Quality (L)*
Perceptual Analyzer Manipulation	*VE relevant priming (P)*	H-P	L-P
	VE irrelevant priming (N)	H-N	L-N

Table 7.1 The four evaluation conditions created by manipulating two activation levels on each input layer of the model.

7.2 Applying the model to predict relative presence levels in the four novel conditions

This section will use the connectionist model to predict the relative presence levels for each of the four novel conditions defined in 7.1.3 above. It is important to note that the connectionist model is not yet sophisticated enough to predict presence quantitatively, but it is possible to predict, for a number of different initial model states, the relative presence levels between them. The examples below make use of the same model used in the chapter 6, which is presented below in Figure 7-1 as a reference.

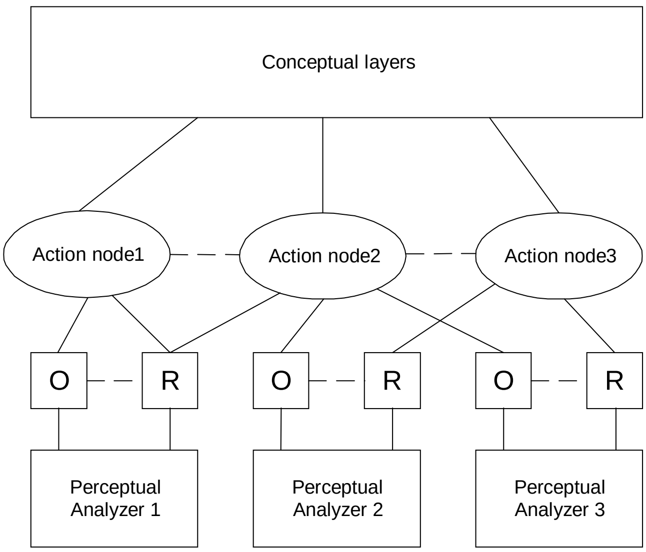


Figure 7-1: The model used in the predictions

As with the previous examples, action node 2 will represent the "presence node" – that is, the action node which encodes the schema appropriate to the VE the subject is operating in. The level of activation of this node at the time the network settles will indicate the level of presence experienced by the subject.

7.2.1 Expressing the conceptual layer manipulations in the model

The *VE relevant priming* manipulation of the conceptual layers requires the conceptual layers to contribute activation to the "presence node" – action node 2. This will be modeled by adding top-down activation to action node 2 by way of the connection between itself and the conceptual layers.

The *VE irrelevant priming* manipulation represents a state of priming where some action node not related to the VE appropriate schema is active. The action layer is organized with action nodes encoding similar schemas physically close together, so the action node activated in this condition will be physically far away from action node 2, and is not likely to exert any influence over action node 2. As this model shows only three action nodes, the activated action node is not indicated. This condition is thus modeled by not including any top-down activation.

7.2.2 Expressing the perceptual analyzer manipulations in the model

The *High stimulus quality* condition represents a condition with many perceptual analyzers stimulated, and the analyzers interpret the stimuli as being objects rather than renditions. This is modeled by activating the O nodes of all three perceptual analyzers included in the network, and allowing no activation in the R nodes.

The *Low stimulus quality* condition represents a condition with a fewer analyzers activated, and the stimuli are such that they are interpreted as renditions rather than as objects. This is modeled by highly activating the R nodes of the perceptual analyzers.

7.2.3 Modeling the *High stimulus quality/VE relevant priming* condition

In this condition, the O nodes of the perceptual analyzers are highly activated, with almost no activation in the R nodes. Due to the priming manipulation, the conceptual layers are contributing activation to action node 2. These initial conditions are presented in Figure 7-2.

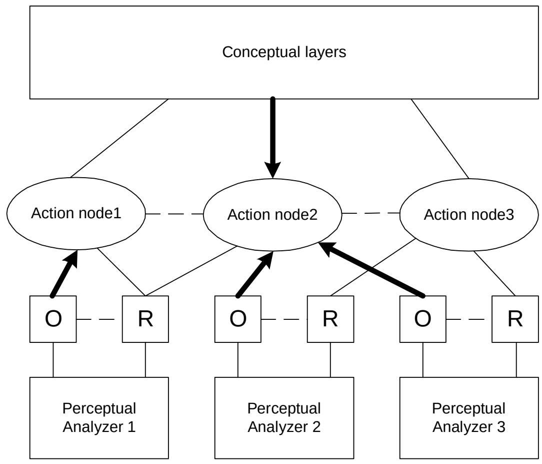


Figure 7-2: Initial model state for the *High stimulus quality/VE relevant priming* condition

Once the model begins processing, action nodes 1 and 2 begin to receive activation. Action node 1 receives activation from the O node of perceptual analyzer 1, while action node 2 receives activation from the O nodes of perceptual analyzers 2 and 3, as well as from the conceptual layers due to the priming manipulation. As these action nodes become activated, they begin to compete and inhibit each other. Action node 2, receiving activation from three sources, will be able to successfully inhibit action node 1, although action node 1 will have inhibited it to some degree as well. The final state, illustrated in Figure 7-3, is that action node 2 receives activation from three sources, and is only inhibited by one, and thus remains highly activated, leading to a high degree of presence.

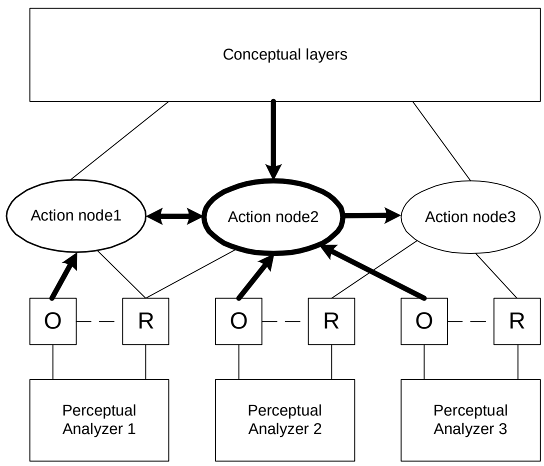


Figure 7-3: Final state for the *High stimulus quality/VE relevant priming* condition

7.2.4 Modeling the *High stimulus quality/VE irrelevant priming* condition

In this condition, all of the perceptual analyzers' O nodes are highly activated, with almost no activation in the R nodes. Due to the priming manipulation, the conceptual layers contribute no activation to action node 2. The initial conditions are presented in Figure 7-4 below.

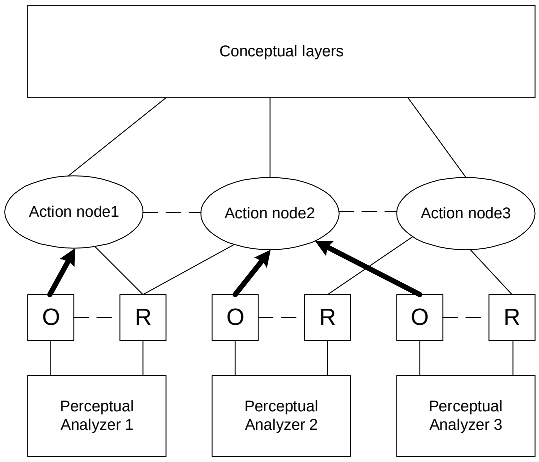


Figure 7-4: Initial model state for the *High stimulus quality/VE irrelevant priming* condition

When the model is started, action nodes 1 and 2 receive bottom-up activation from the perceptual analyzers. Action node 1 receives activation from the O node of perceptual analyzer 1 and action node 2 receives activation from the O nodes of perceptual analyzers 2 and 3. As the activation increases in the action nodes, they begin to compete through the inhibitory connection between them. Action node 2 is more activated than action node 1 (it is receiving activation from two sources, while action node 1 only receives activation from one source), and will tend to inhibit its neighbour more vigorously. When the model settles, action node 2 remains the most active node in the action layer, as it is receiving activation from two sources and inhibition from one. Action node 1 on the other hand, is receiving activation only from one source, and inhibition from one strong source. The final state is shown in Figure 7-5.

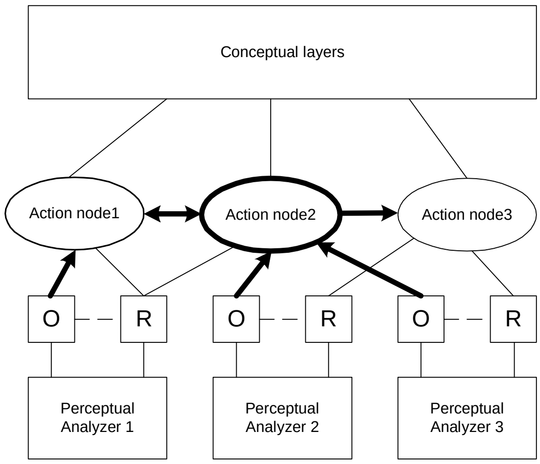


Figure 7-5: Final state for the *High stimulus quality/VE irrelevant priming* condition

7.2.5 Modeling the *Low stimulus quality/VE relevant priming* condition

Unlike the previous two conditions described, this condition has low stimulus quality, and following the conventions outlined in 7.2.2 above, the R node of each perceptual analyzer will be highly activated, while the O nodes will begin with no activation. Due to the conceptual layer manipulation, action node 2 will receive considerable top-down activation. This state is represented in Figure 7-6 below.

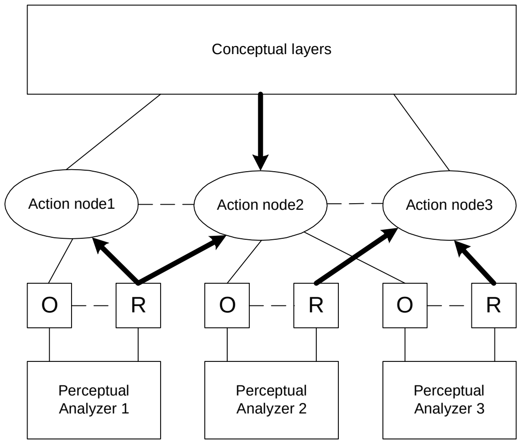


Figure 7-6: Initial model state for the *Low stimulus quality*/*VE relevant priming* condition

When the model is run, the activation from the R nodes beings to spread upwards, activating action nodes 1, 2 and 3. Action node 3 receives activation from the R nodes of perceptual analyzers 2 and 3, and thus becomes significantly more activated than either action node 1 or 2, which share the activation of the R node of perceptual analyzer 1. Simultaneously, action node 2 receives top-down activation from the conceptual layers as a consequence of the priming manipulation. As the action nodes become activated, they begin to compete via the inhibitory connections. Action node 2 (the “presence node” in this example) which is receiving input from two sources (the conceptual layer and the R node of perceptual analyzer 1) will inhibit its neighbours strongly. Thus action node 1 (which receives only activation from the R node of perceptual analyzer 1) will be overwhelmed by the inhibitory activation from action node 2. Action node 3 also receives activation from two sources (the R nodes of perceptual analyzers 2 and 3), so it will resist the inhibition of action node 2. The final, stable state of the model leaves action nodes 2 and 3 with similar levels of activation. This is shown in Figure 7-7 below.

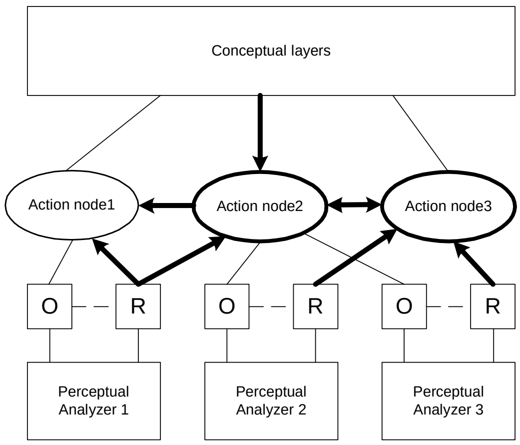


Figure 7-7: Final state for the *Low stimulus quality*/*VE relevant priming* condition

7.2.6 Modeling the Low stimulus quality/VE irrelevant priming condition

Similar to the condition described in 7.2.5 above, this condition has low stimulus quality, and thus the R node of each perceptual analyzer will be highly activated, while the O nodes will begin with almost no activation. As a consequence of the manipulation of the conceptual layers, the action nodes receiving top-down activation will be far from the action node 2 and will thus have a negligible effect on its activation. This initial state is illustrated in Figure 7-8 below.

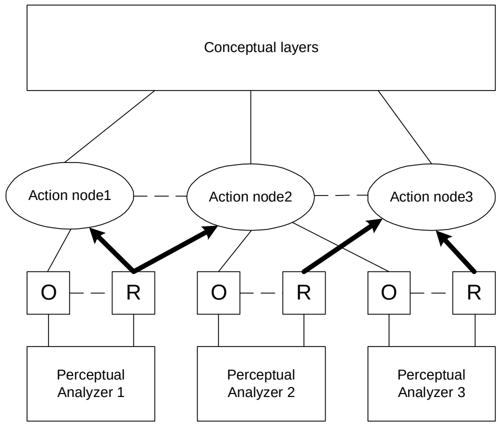


Figure 7-8: Initial model state for the *Low stimulus quality*/*VE irrelevant priming* condition

As the activation spreads, action nodes 1 and 2 receive activation from the R node of perceptual analyzer 1, and begin to compete. As each receives the same amount of bottom-up activation, the competition between them leads to a stalemate. Action node 3, however, receives activation from 2 sources – the R nodes of perceptual analyzer 2 and 3. This allows it to effectively inhibit action node 2. The resulting state contains action node 3 as the most activated of the model, with action node 2 having only a slight level of activation. This final state is represented in Figure 7-9 below.

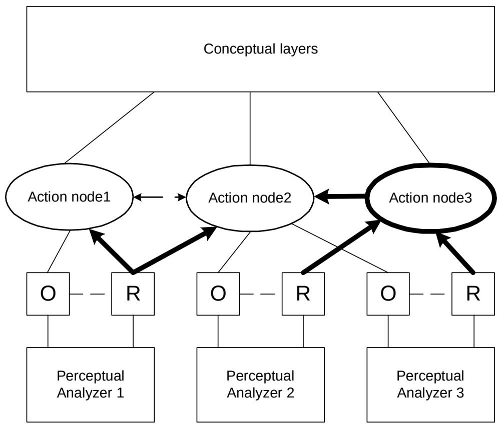


Figure 7-9: Final model state for the *Low stimulus quality*/*VE irrelevant priming* condition

7.3 The predicted relative level of presence in each of the four novel conditions

As can be seen from the application of the model to each of the four novel conditions in 7.2 above, the interaction of conceptual layer manipulation and stimulus quality is predicted to have a great impact on the level of presence experienced by a subject. In the case of the *High stimulus quality* conditions, the model predicts that the convergence of top-down (conceptual layer) activation and bottom-up (perceptual analyzer) activation will lead to a significant increase in presence, but in the case of the *Low stimulus quality* conditions, the contribution of top-down activation seems negligible, as both priming conditions conclude with a low level of action in the "presence node" (action node 2). These predictions are summarized in
Table 7.2, and presented graphically in Figure 7-10 below.

		Conceptual Layer Manipulation	
		High Stimulus Quality (H)	*Low Stimulus Quality (L)*
Perceptual Analyzer Manipulation	*VE relevant priming (P)*	High	Low
	VE irrelevant priming (N)	Moderate	Low

Table 7.2: Relative presence levels predicted by the model for each of the four novel conditions created by manipulating two activation levels on each input layer of the model

7.3.1 Specific results expected

When empirically testing these results, the following patterns should be evident in the data, according to the predictions made by the model:

1. An overall difference in presence between the *High stimulus quality* and the *Low stimulus quality* conditions, with the *High stimulus quality* condition producing higher presence scores. This is can be better appreciated in Table 7.1; note that the *High Stimulus Quality* column contains higher predicted presence values that those in the *Low Stimulus Quality* column.

2. A difference between the *High stimulus quality/VE relevant priming* and the *High stimulus quality/VE irrelevant priming*, with the relevant priming leading to higher presence scores. In
3. Table **7.2**, this difference exists between the two values in the *High Stimulus Quality* column.

4. Little or no difference between the *Low stimulus quality/VE relevant priming* and the *Low stimulus quality/VE irrelevant priming* conditions. Again,
5. Table **7.2** shows this; note the *Low Stimulus Quality* column. Both predictions are for the a low presence level.

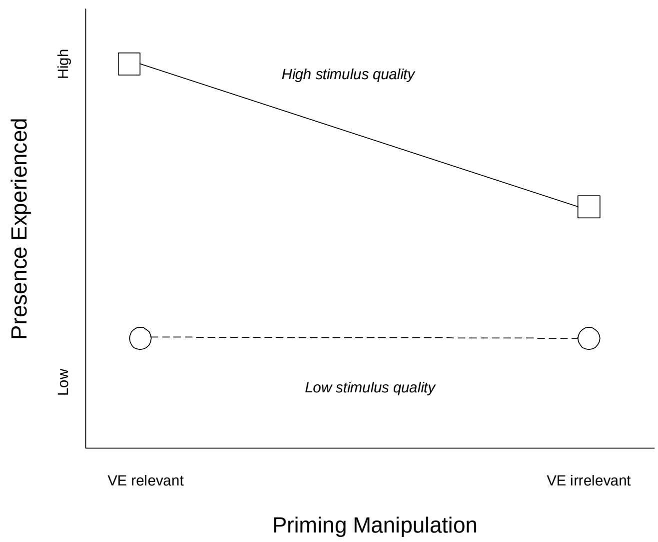


Figure 7-10: Predicted means of relative presence experienced by subjects in each of the four test conditions

7.4 Testing the independence of O/R node activation and presence

Examination of our model proposed here reveals that a state of *presence* in any one environment can be arrived at by several ways. For example, in the models used in these predictions (as in Figure 7-1), one way for the perceptual layers to activate action node 2 (the node we use to represent presence in our examples) is by means of the O nodes of perceptual analyzers 2 and 3. In practical terms, this implies that to increase presence, one must increase the quality (and by implication, the realism) of the display. However, the description of the model's architecture allows for R nodes to activate action nodes as well (for example, in the case of the R node connection between perceptual analyzer 1 and action node 2). Because a path exists between the perceptual analyzer and the action layer which does not include an O node, this implies that presence can be independent of O node activation, and that it is possible for presence to occur as a consequence of R node activation only (even if this means a low level of display realism).

To evaluate this idea, it is necessary to create a display of a virtual environment that consists mostly of renditions of objects rather than objects themselves. Possible ways of doing this include using still images, or text descriptions of rooms, as used in Multi-User Dungeons (MUDs – see Towell & Towell, 1997, for an investigation into presence in virtual environments displayed using text only). Once this display has been created, a comparison can be made of the presence levels experienced by subjects viewing it with the presence levels experienced by subjects viewing other display types.

If it is possible to create using presence using the activation of R nodes only, then it is a methodological requirement that the "R node only" display produce presence levels comparable to those produced by a display which activates a combination of O and R nodes. This is necessary because there exists no absolute method for measuring the activation level of a particular set of action nodes. Thus, the presence level produced by the "R node only" display might in fact be an effective "zero presence" level. However, if the presence level is comparable to a display which stimulates both O and R nodes, then it is easier to sustain the argument that "R node only" displays can produce presence levels similar to those of displays with higher image fidelity.

To test this idea empirically, we propose comparing three displays of a virtual environment. The first makes use of high quality stimuli, presented to several modalities, which increases the probability that a number of O nodes will be activated. The second display uses lower quality presented to several modalities, increasing the probability of activating a combination of O and R nodes. The final system uses text descriptions of the VE, ensuring that only R nodes are activated. This effectively creates three conditions – one where mostly O nodes are activated, one where a combination or O and R nodes are activated, and one where mostly R nodes are activated.

7.4.1 Specific results expected

Based on the arguments presented above, if presence is indeed independent of O node activation, the following results are expected. These are presented graphically in Figure 7-11 below.

1. Subjects viewing the text only display will experience less presence than those viewing the high stimulus quality display

2. Subjects viewing the text only display will experience the same level of presence than those viewing the low stimulus quality display

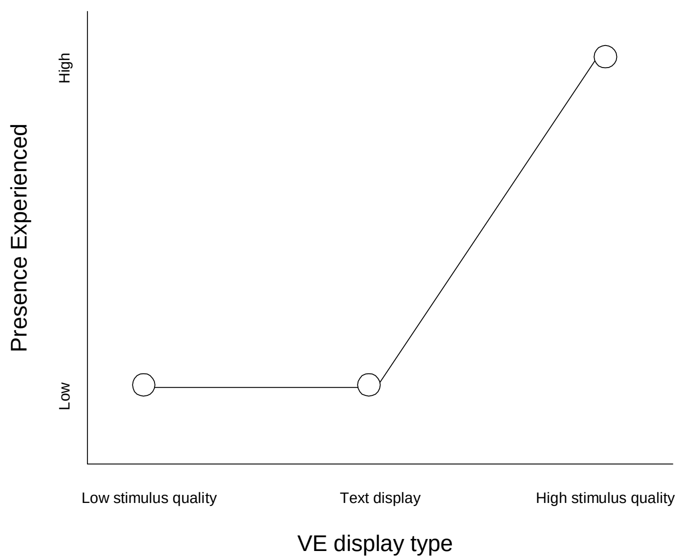


Figure 7-11: Predicted means of relative presence experienced by subjects viewing the a VE with a low quality graphical display, a text display and a high quality text display

Chapter 8

Experiment 1

8.1 Priming and Stimulus quality: Hypotheses

This experiment sets out to falsify the main tenets of the presence model set out in Chapter five directly. These are:

(a) The quality of the VR stimuli have an effect on presence

(b) Priming participants before they enter a virtual environment can have an effect on presence

(c) Presence levels will depend on the interaction of priming and stimulus quality.

The first hypothesis follows directly from the mounting body of evidence which suggests that virtual environments which are of higher realism and which appeal to a greater number of modalities will lead to greater levels of presence. To test this idea, we have implemented two virtual environments each in two levels of quality, and each participant is placed into one of these two conditions. In the *low stimulus quality condition*, the environment is not textured, Lambertian lighting is used, and no sound is present. In the *high stimulus quality condition*, the same geometry is used as in the low quality condition, but in this condition, texturing is used, as is radiosity. There is also directional and environmental sound. The two conditions vary not only in realism, but also in the number of modalities they appeal to. The hypothesis then is that *the high stimulus quality condition will lead to higher presence levels than will the low stimulus quality condition*

The second hypothesis is based on the constructive perception thesis, which states that perceptions of ambiguous stimuli are more accurately processed by a participant which has been primed for that task. We believe that higher presence levels will result if the participant has been primed before exploring the virtual environment, because the priming will help to resolve ambiguities and inconsistencies in the environment, leading to a more coherent experience, and thus higher presence. To test this idea, we place participants into one of two conditions. In the *relevant priming* condition, participants were asked to read a booklet which was related in theme to the virtual environment (for example, participants who were to enter a hospital virtual environment were asked to read a booklet about emergency room procedures). In the *irrelevant priming* condition, subjects were asked to read a booklet which was not related in theme to the virtual environment. The hypothesis then is that *the priming relevant condition will lead to higher presence levels than will the priming not relevant condition.*

The third hypothesis is based on the proposed effects of lateral inhibition in the action layer of the environmental perception connectionist network presented in this document. The effect of these later connections is that if nodes in the action layer are highly stimulated, they will vigorously inhibit their neighbours in the same layer, and that a node which has been thus inhibited is less likely to contribute to the sense of presence for that environment. We expect that the level of presence will vary for a given stimulus quality level if the priming is altered. To test this idea we will examine the interaction effects of priming and stimulus quality on presence. The hypothesis is thus that *presence levels will vary as a function of the interaction between priming and stimulus quality.*

8.2 Variables

The main variables in this experiment are *presence*, *stimulus quality* and *priming.* Secondary variables used are *immersive tendencies* and *fun experienced in the virtual environment*.

8.2.1 Definition of variables

- *Presence:* this is a complex concept, and is discussed in depth in Chapter two. Presence is used as a dependent variable in this experiment.

- *Stimulus quality:* this is the quality with which the virtual environment is rendered. It depends on how closely the virtual rendering matches the virtual environment. This can be understood in several ways. One way of understanding this concept is as photo-realism. The higher the stimulus quality, the more photo realistic the rendering, although stimulus quality also includes the notion of multi modality ("photo-realism" for sound, etc as well as for images). Another way of understanding stimulus quality is through the notion of information load. A rendering of an environment which carries more environmentally relevant information has higher stimulus quality that one which carries less environmentally relevant information.

- *Priming:* this is the notion of providing the participant with information which is relevant to the virtual environment before they experience the environment. The idea of priming has existed in cognitive psychology in the constructive perception school as a way of influencing identification of ambiguous stimuli for many years, and is quite well understood.

- *Immersive tendencies :* this variable is defined by Witmer and Singer (1990). This is the capacity of a person to become immersed, be it in VR, cinema or any other immersive medium. Being a property of a person rather than a system, it is independent of any particular display device or medium, but some studies show that it is a predictor of presence. Its measurement is based on the subject's previous experiences with immersive media.

- *Fun experienced in the virtual environment:* this is the enjoyment experienced by the participant during the time of their immersion in the virtual environment. It is based on the subject's memory of their experience in the virtual environment.

8.2.2 Operationalization of variables.

- *Presence:* three measures of presence were used, namely the Presence Questionnaire of Witmer & Singer, the questionnaire of Slater, Usoh & Steed and the Contents of Consciousness Inventory introduced in this document. Presence is used as a dependent variable.

  The Presence Questionnaire (PQ) consists of 32 semantic differential items, each giving a score in the range from 1 to 7. In 28 of the items, a score of 1 indicates a response consistent with low presence, while a score of 7 indicates a score consistent with high presence. The remaining four PQ items (items 19, 22, 23 and 26) are reversed, with a score of 1 consistent with high presence, and a score of 7 being consistent with low presence. For the sake of simplicity, after administering the PQ, we reversed the response of these four items so that a higher total PQ score can always be understood as indicating more presence. A subject's presence score on the PQ is simply the sum of their responses on the 32 items. A completed PQ yields a score from 32 to 224. The PQ appears in section G.2 of Appendix G.

The Slater, Usoh & Steed questionnaire (SUS) consists of 6 semantic differential items, each giving a score in the range from 1 to 7. For each item, a score of 1 indicates a response consistent with low presence, while a score of 7 indicates a response consistent with high presence. A subject's presence score on the SUS is simply the sum of their responses on the 6 items. A complete SUS produces a score of 6 to 42. The SUS appears in section G.1 of Appendix G.

The Contents of Consciousness Inventory (COCI) consists of 10 forced choice items, each giving a score of 1 or 0. A score of 1 indicates that the participant has selected the option which matches the theme of the virtual environment, and 0 indicates that the subject has not selected the option that matches the theme of the virtual environment. The COCI score is simply the sum of all COCI items. A complete COCI gives a score ranging from 0 to 10. The COCI implementation is described in section A.6 of Appendix A. The COCI items used are listed in Appendix F.

- *Stimulus Quality:* In this experiment, stimulus quality is an independent variable, manipulated into two levels: 'l' or low stimulus quality and 'h' high stimulus quality. In both levels, the same virtual environment and geometry is presented. In the low stimulus quality level (the 'l' level), the environment is not textured, Lambertian lighting is used, and no sound is present. In the high stimulus (the 'h' level), radiosity is used and the geometry is textured. There is also sound, both environmental (sounds which emanate from no particular place and serve as an aural backdrop, such as the sound of thunder) and positional sound (sounds originating from a particular place, such as the sound of ringing from a telephone). Appendices C and D contain images illustrating the visual differences between the two levels of stimulus quality.

- *Priming:* priming is an independent variable in this experiment and is therefore manipulated rather than measured. Two levels of priming exist: a relevant priming condition (abbreviated 'p') and an irrelevant priming condition (abbreviated 'n'). In the priming relevant condition ('p' level), the participant is allowed to examine, prior to the VR experience, a booklet which is related in theme to the virtual environment they are about to experience. In the irrelevant priming condition ('n' level), the same manipulation occurs, but the booklet is not related to the virtual environment they are about to experience. Details of the contents of the booklets are stated in section 8.4.6. The booklets themselves are reproduced in Appendix H.

- *Immersive tendencies:* this variable is measured by means of the Immersive Tendencies Questionnaire (ITQ) compiled by Witmer & Singer (1990). The ITQ consists of 32 semantic differential items, each giving a score in the range 1 to 7, with a response of 1 being consistent with low immersive tendencies, and a score of 7 being consistent with an individual who tends to be highly immersive. A participant's ITQ score is simply the sum of all their responses. A complete ITQ provides a score which ranges between 32 and 224. The ITQ is reproduced in section G.3 of Appendix G.

- *Fun experienced in the environment:* this variable is measured with a single Likert-type item, in the range 1 to 10. A score of 1 is consistent with having felt no fun, while a score of 10 is consistent with having had a lot of fun. A participant's score is simply their response to this item, in the range 1 to 10. This item is reproduced in sections G.4 and G.5 of Appendix G.

8.3 Design

This experiment makes use of a 2x2x2 (stimulus quality x priming x virtual environment) factorial design, the final factor being a within-subjects factor. This creates 8 cells in the design, and each participant contributes data to two of the cells. This type of design demands a large number of participants, but provides the opportunity to investigate interactions between factors.

8.4 Materials

8.4.1 Venue

The experiment was conducted in a dedicated room in which the lighting and extraneous noise could be controlled by means of the door and window blinds. Inside this room, three computers were placed so that up to three participants could be run simultaneously. Partitions were placed between each computer so that participants could not see each other or their displays (although they entered the room together and were aware that others were in the room with them). The experimenter could stand behind a partition on one end of the room and observe the participants without them being distracted by his observation.

8.4.2 Computers

3 computers were used for this experiment, to allow 3 participants to be run simultaneously. All computers were equipped with the same hardware and software, namely

- AMD Athlon 700 MHz processor
- 128 MB RAM
- GeForce 2 MX 32MB graphics card
- 17” monitor displaying a 640x480x16 graphical stream at an average of 15Hz
- Windows 2000
- Sound Blaster sound card
- Stereo headphones
- Keyboard and mouse
- DAVE v1.0 (cf. Appendix A for a description)

The computers were not connected to a network during the experiment.

8.4.3 VR display system

This experiment used our DAVE v1.0 tool to display the virtual environments. This tool is in more detail in Appendix A. This system makes use of a desktop VR system, which render via means of a CRT monitor and stereo headphones only. Input was by means of keyboard and mouse.

8.4.4 Participants' task

The participants were given a simple task to perform during their exposure in the VE. Inside the VEs, a series of small boxes, called tokens, were placed by the experimenter. Only one of the tokens was visible at any one time. The participant's task was to search for these tokens and collect them by colliding with the token. When a token had been found, it disappeared, and the next token in the series would appear in another location. To help the participants find the tokens, a "token scanner" was shown on the bottom-right hand corner of their display. This scanner showed a participant their position relative to the current token. A circle of 8 lights showed the relative angle to the token, while a set of two lights, one above the other, showed if the token was above or below the present position. The participants could also query the scanner for the distance to the token by clicking the left mouse button. When clicked, the scanner would produce a sound whose pitch indicated distance to the token – a high pitch sound if the token was nearby, and a low pitch sound if the token was further away. Two versions of the scanner were created, to better fit into the themes of the VEs. These differed only in their appearance, as shown in Figure 8-1. The "modern scanner" was intended for use in the hospital VE, and the "old scanner" was used in the monastery VE.

Figure 8-1: The token scanners. On the left is the "modern" version used in the hospital VEs, and on the right is the "old" version used in the monastery VEs. Both show the token to the left of and above the participant.

The purpose of this task was simply to keep the participant's attention focused on the VE, and to ensure that the participants did not stay in a small area of the VE. By placing a token in each major area of the VE, the experimenter could ensure that most participants visited most areas of the VE, without fear of them becoming lost or wandering in circles. No measurements were taken with regard to the task.

8.4.5 Virtual environments

6 DAVE virtual environments were used during the experiment:

- train.cfg – This environment is abstract, and does not represent any particular theme. It is a brick building with 12 sparsely furnished rooms in three levels, and is used to train the participants in the use of the DAVE tool. It features detailed textures, radiosity as well as positional and environmental sounds. A set of 13 tokens is present, with the "modern scanner". No COCI items are defined. Images from this VE can be seen in Appendix E.

- train_coc.cfg – This environment is used to teach the participants about the COCI. It is the same as train.cfg except that the token scanner is the "old scanner" and a set of 6 COCI items is defined. Images from this VE can be seen in Appendix E, and the COCI items used are listed in section F.3 of Appendix F.

- mon_hi.cfg – The monastery, high stimulus quality. This environment represents a medieval monastery and chapel in the countryside. The environment features 3 levels, with 18 furnished rooms. There are detailed textures, radiosity, as well as positional and environmental sound. 10 COCI items are defined. There are 32 tokens, with the “old scanner”. Images from this VE can be seen in Appendix D. The COCI items used are listed in section F.2 of Appendix F.

- mon_lo.cfg – The monastery, low stimulus quality. This environment is the same as the mon_hi.cfg environment, except that there is no sound, the textures are flat colours, and there is no radiosity. Images from this VE can be seen in Appendix D. The COCI items used are listed in section F.2 of Appendix F.

- hosp_hi.cfg – The hospital, high stimulus quality. This environment represents a contemporary hospital in a city. This environment features 3 levels with 21 furnished rooms. There are detailed textures, radiosity, as well as positional and environmental sound. 10 COCI items are defined. There are 36 tokens, with the “modern scanner”. Images from this VE can be seen in Appendix C. The COCI items used are listed in section F.1 of Appendix F.

- hosp_lo.cfg – The hospital, low stimulus quality. This environment is the same as the hosp_hi.cfg environment, except that there is no sound, the textures are flat colours, and there is no radiosity. Images from this VE can be seen in Appendix C. The COCI items used are listed in section F.1 of Appendix F.

8.4.6 Priming materials

Three priming booklets were created; two related in theme to the monastery and hospital virtual environments, and a neutral booklet with a theme unrelated to either virtual environment. Each booklet contained a piece of text, approximately 1000 words long, as well as approximately 4 pictures. In the case of the monastery and hospital themed booklets, the pictures were printed in colour (the neutral booklet’s pictures were not printed in colour as a cost saving measure). The booklets have been reproduced as an appendix to this document.

- *Monastery booklet* - This booklet contained a brief history of the monastic movement in medieval times, and a brief synopsis of everyday life in a monastery. The text was assembled from several web pages on the subject, and consisted of 963 words, and five pictures. Two of the pictures were of monasteries in England and France, and the other three were scans of illuminated religious scripts from the medieval period. This is reproduced in section H.1 of Appendix H.

- *Hospital booklet* – This booklet contained an explanation of a hospital emergency room and how a patient is treated upon entering the emergency room, and was assembled from several web pages on the subject. The booklet consisted of 1022 words and 4 pictures. The pictures were all photographs showing medical equipment, people being treated, and a surgeon performing an operation. This is reproduced in section H.2 of Appendix H.

- *Neutral booklet* – This booklet contained a description of driving a steam train. It explains the procedures of starting the train, as well as some personal anecdotes by train drivers, and was assembled from several web pages on the subject. The booklet consisted of 1051 words and 3 pictures. The pictures showed steam trains, as well as the controls of a locomotive. This is reproduced in section H.3 of Appendix H.

8.4.7 Measurement instruments

Four separate sets of scales were administered to each participant. These included 2 separate COCI sets (one themed to the monastery, the other to the hospital), which were administered electronically during the exploration stage. After each of each of the exploration stages, a set of questionnaires printed on paper were administered. The questionnaire booklets, as presented to the participants, have been reproduced as an appendix to this document.

The first questionnaire series included the following scales, in this order:

- Slater, Usoh & Steed (2000) scale
- Presence Questionnaire (Witmer & Singer, 1998)
- Form 100 (developed for this experiment)

The second questionnaire series included the following scales, in this order:

- Slater, Usoh & Steed (2000) scale
- Presence Questionnaire (Witmer & Singer, 1998)
- Form 200 (developed for this experiment)
- Immersive Tendencies Questionnaire (Witmer & Singer, 1998)

8.4.8 Experimental schedule

In order to prevent repetition and associated learning effects on any of the 3 factors, each participant explored 2 virtual environments. In this way, there was no need to repeat environment, priming condition or stimulus quality. For instance, if a participant explored the monastery in high stimulus quality with priming in the first exploration stage, then the second exploration stage would see that participant exploring the hospital in low stimulus quality with no priming. This means that each participant contributes data to two of the eight cells of the design.

8.5 Participants

Posters advertising the experiment were placed around the Computer Science building at the University of Cape Town. This building is used not only by computer science students, but also by undergraduates of the Science faculty, the Commerce faculty and the Health Sciences faculty. The posters asked for participation in a virtual reality experiment for a payment of 20 Rand. The posters did not express preference for any particular group. Volunteers were asked to write their name and contact telephone number on a sign-up sheet outside the experimental venue. The sign-up sheets contained details of the times available for participation.

Demographic details of the participants were not recorded; however, all were volunteers. Although details were not recorded, the group included both men and women (with a much higher proportion of men) of various ethnic groups. Almost all participants were in their early twenties. The level of familiarity with computers seemed to vary greatly, from one participant who stated he knew nothing about computers, to another who after the experiment engaged the experimenter in a discussion of the technical details of the DAVE tool.

A total of 60 volunteers wrote down their name and contact telephone number on the sign up sheet. Of those, 55 actually arrived at the agreed time to participate in the experiment. The others did not appear at the experimental venue. This constitutes a 91.7% response rate.

8.6 Procedure

8.6.1 Instruction and Training stage

The experiment was run in a dedicated room so that lighting and noise could be controlled to reduce distractions. The room contained three computers with partitions between each. The participants were asked to sit down at a computer of their choice. Their selection of computer placed them in an experimental condition, as a random schedule based on computer had been worked out previously. The basic purpose of the experiment was explained to them. They were told that the experiment wanted to look at a person's thought process while visiting a virtual environment differed to their thought process while visiting a real place. They were then given a basic instruction of their task; namely that they were to be tourists in the virtual environment, and that their main task should be to take in the sights and sounds of the virtual environment. The DAVE tool was then started with the train.cfg virtual environment (which included tokens but not the COCI items), and the basic movement interface was explained to them. They were allowed to practice moving under the supervision of the experimenter, who helped those participants who were having problems. The participants were allowed to practice until the experimenter was satisfied that all the participants understood the principles of the movement interface. This usually took no longer than 3 minutes.

At this point, the participants were introduced to the idea of the tokens and the scanner. They were told that in order to help them explore the virtual environment more thoroughly, a series of tokens were scattered in the virtual environment, and that if one went about collecting them, this would lead one to all the interesting places of the environment. They were shown the workings of the scanner, and allowed to practice finding tokens in the training environment under the supervision of the experimenter. Once the experiment was satisfied that all the participants understood the process (this usually took no longer than 3 minutes), he reminded the participants that collecting tokens was only a secondary task, and that their main task should be to explore the environment (this was done to avoid the participants from focusing on the token collecting task at the expense of taking in the virtual environment). The DAVE tool was shut down, and the participants were explained the procedure of the COCI. They were told that the purpose of the COCI was to assess their thought process, and that it would occur periodically at intervals of about 1 minute. They were shown what an item looked like, and were shown the controls used to respond. Once the explanation was complete, the DAVE tool was started using the train_coc.cfg virtual environment (which includes tokens as well as a set of COCI items). The participants were asked to begin exploring until the first item appeared. Once the item appeared, the experimenter explained the process again, emphasizing the point that none of the four options were "correct" but that the participants should choose the one they thought fit the word fragment best. The participants were then allowed to practice under the supervision of the experimenter, who ensured that all the participants understood all the interface aspects correctly, and offered correction where necessary. This practice session was allowed to run for 5 to 6 minutes. The DAVE tool was then shut down.

8.6.2 Priming stage

The participants were then told that the training was over and that the experiment was about to begin. The basic procedure was explained to them, namely that they would be given a booklet to read, followed by exploring a virtual environment, followed by filling out of questionnaires. Before they were given the priming materials, the participants were given instructions on how to read it. They were told that the booklet was long, but that only 5 minutes would be given to read it. It was emphasized that it was not important to finish the entire booklet in the time given, but rather that they should read slowly, carefully examine the pictures, and think about the things written in the text. The door of the room was closed, and the participants were then given the priming material as dictated by the schedule. While the participants read, each computer was prepared with the virtual environment as per the schedule. After 5 minutes, the priming materials were taken away.

8.6.3 Exploration stage

The room's lights turned off, and the DAVE tool started. The subjects were instructed to begin exploring the virtual environment. The experimenter remained in the room, but observed the participants from a covert position so that the participants were not distracted. Once all of the participants had completed the entire set of COCI items (which took between 11 and 15 minutes), the exploration stage was concluded.

8.6.4 Questionnaire stage

The room's lights were turned on, the DAVE tool shut down, and the participants were handed the first series of questionnaires to complete. The participants were then given time to complete the entire set of questions, which usually took between 10 and 15 minutes.

8.6.5 Second iteration

Once the questionnaires were complete, the participants were told that they were to explore one more virtual environment. The same basic procedure as above was repeated, from the priming stage till the questionnaire stage. At the beginning of this second priming stage, the participants were again reminded of the importance of not rushing through the booklet, but rather reading carefully. During the second questionnaire stage, the second series of questionnaires were administered. This slightly longer set of questionnaires took between 15 and 20 minutes to complete.

8.6.6 Completion and preparation stage

Once the participants had completed the questionnaires, they were asked to sign a receipt, and were paid 20 Rand for participating. When the participants had left, the experimenter checked the data output file in each computer for completeness, and prepared each computer for the next set of participants. The paper questionnaires were then coded and stored.

8.7 Analysis of results

In this section we present an analysis of the data collected in this experiment. We first present descriptive statistics for each of the variables, followed by an inferential analysis. The analyses presented were conducted using Microsoft Excel 2000 and StatSoft Statistica 5.5.

8.7.1 Categorization of participants into conditions

Most of the 55 subjects contributed 2 sets of observations (some did not due to time restraints, equipment failures and other such unfortunate occurrences). Table 8.1 shows the frequency of observations in each of the four cells of the design (the virtual environment condition was collapsed).

	Priming		
Stimulus Quality	Relevant priming (‘p’)	Irrelevant priming (‘n’)	*Row totals*
Low quality (‘l’)	24	28	*52*
High Quality (‘h’)	27	24	*51*
Column totals	*51*	*52*	**103 (Grand Total)**

Table 8.1: Number of observations in each of the conditions

8.7.2 Descriptive statistics for COCI, SUS, FUN, PQ and ITQ

Each of the 103 observations was made on Witmer and Singer’s Presence Questionnaire (PQ), Slater, Usoh & Steed’s questionnaire (SUS) and the Contents of Consciousness Inventory (COCI). Furthermore, in the Form 100 and Form 200 questionnaires, participants were asked to rate their sense of fun on a scale of 1-10 (FUN). 44 of the 55 participants also completed Witmer and Singer’s Immersive Tendencies Questionnaire (ITQ). The descriptive statistics for these variables are presented in Table 8.2.

Variable	Valid N	Mean	-95% CI	+95% CI	Min	Max	Std.Dev.
COCI	103	5.0777	4.5458	5.6095	0	10	2.72134
SUS	103	25.0291	23.4853	26.5730	0	39	7.89944
FUN	102	6.2157	5.7388	6.6925	1	10	2.42768
PQ	103	151.6019	146.2269	156.9770	95	252	27.50235
ITQ	44	155.1136	149.1902	161.0370	104	190	19.48311

Table 8.2: Descriptive statistics for COCI, SUS, FUN, PQ and ITQ

8.7.3 Correlations between COCI, SUS, FUN and PQ

To test if the various presence measurements are valid, we conducted a series of correlations to test the concurrent validity of these measures. Also, we looked for correlations between FUN and the various presence measures. A correlation matrix (n=43) was computed from the variables FUN, PQ, ITQ, COCI and SUS. Table 8.3 presents a summary of the results, with significant correlations ($p < 0.05$) being in bold.

	COCI	SUS	FUN	PQ	ITQ
COCI	1.0	*.01*	**0.33**	*.24*	*.05*
SUS	*.01*	1.0	**.57**	**.76**	*-.02*
FUN	**.33**	**.57**	1.0	**.74**	*.04*
PQ	*.24*	**.76**	**.74**	1.0	*.08*
ITQ	*.05*	*-.02*	*.04*	*.08*	1.0

Table 8.3: Correlations between COCI, SUS, FUN, PQ and ITQ. Significant correlations ($p < 0.05$) in bold

8.7.4 Factorial ANOVA with Stimulus Quality and Priming as independent variables and SUS, PQ and COCI as dependent variables

To test the effect of stimulus quality and priming on the measures of presence, 2x2 factorial ANOVAs were conducted. The data used were the combination from the monastery and hospital environments. This same analysis was conducted using SUS, PQ and COCI as dependent variables.

COCI as the dependent variable – all effects: An investigation of the effects shows that there is no significant interaction between stimulus quality and priming ($F(1,99) = 0.26$ $p > 0.611$). There is also no significant main effect for stimulus quality ($F(1,99) = 1.89$ $p > 0.17$). However, there does exist a significant main effect for priming ($F(1,99) = 8.46$ $p < 0.005$). A summary of these statistics is presented in Table 8.4. A protected t-test of the significant effect (priming) reveals a t value of 3.011 (df = 101, $p < 0.004$). A means plot for this effect is presented in Figure 8-2.

	df Effect	MS Effect	df Error	MS Error	F	p-level
Stim Qual	1	13.01460	99	6.850896	1.899693	.171220
Priming	**1**	**57.99200**	**99**	**6.850896**	**8.464877**	**.004471**
Interaction	1	1.78013	99	6.850896	.259840	.611365

Table 8.4: Summary table of effects for a 2x2 factorial ANOVA with COCI as the D.V. Significant effects ($p < 0.05$) marked in bold.

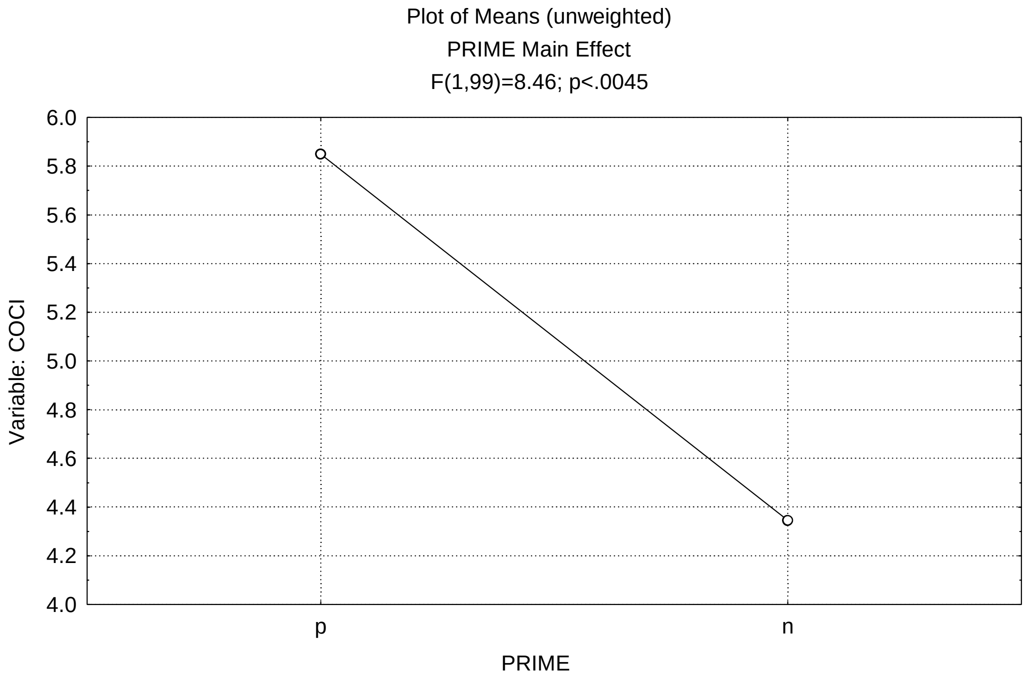


Figure 8-2: Means plot of the main effect of priming on COCI

<u>*SUS as the dependent variable – all effects:*</u> This variable presents a significant interaction between stimulus quality and priming (F(1, 99) = 10.18 p < 0.003). The means plot of this effect is shown in Figure 8-4. There is also a significant main effect on stimulus quality (F(1,99) = 9.64 p < 0.002). The means plot of this effect is shown in Figure 8-3. There is, however, no significant main effect in priming (F(1,99) = 0.17 p > 0.65). This effect information is summarized in Table 8.5.

	df Effect	MS Effect	df Error	MS Error	F	p-level
Stim Qual	**1**	**516.8157**	**99**	**53.61860**	**9.63874**	**.002485**
Priming	1	8.9994	99	53.61860	.16784	.682923
Interaction	**1**	**545.9874**	**99**	**53.61860**	**10.18280**	**.001900**

Table 8.5: Summary table of effects for a 2x2 factorial ANOVA with SUS as the D.V. Significant effects (p < 0.05) marked in bold.

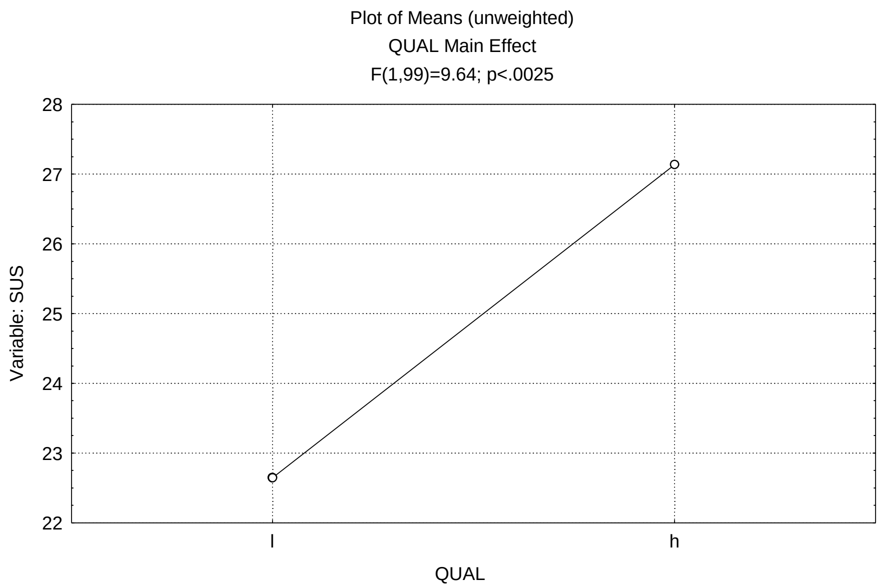


Figure 8-3: Means plot of the main effect of stimulus quality on SUS

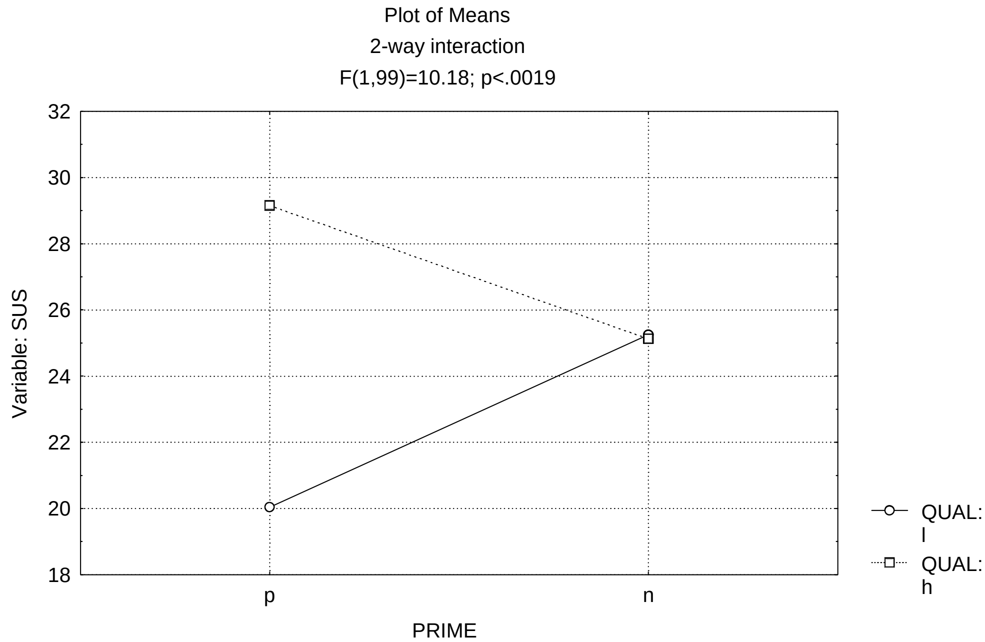


Figure 8-4: Means plot of the interaction between stimulus quality and priming on SUS

PQ as the dependent variable – all effects: This variable behaves in a similar way to SUS, which comes as little surprise as they are highly correlated. The interaction between stimulus quality and priming (shown in Figure 8-5) is significant ($F(1,99) = 4.23$ $p < 0.05$), as is the main effect of stimulus quality on PQ shown on Figure 8-6 ($F(1,99) = 5.99$ $p < 0.02$). The main effect of priming on PQ is not significant ($F(1,99) = 0.23$ $p > 0.63$). Table 8.6 presents a summary of the effects.

	df Effect	MS Effect	Df Error	MS Error	F	p-level
Stim Qual	**1**	**4221.079**	**99**	**704.3596**	**5.992791**	**.016125**
Priming	1	160.254	99	704.3596	.227518	.634422
Interaction	**1**	**2981.610**	**99**	**704.3596**	**4.233080**	**.042272**

Table 8.6: Summary table of effects for a 2x2 factorial ANOVA with PQ as the D.V. Significant effects ($p < 0.05$) marked in bold.

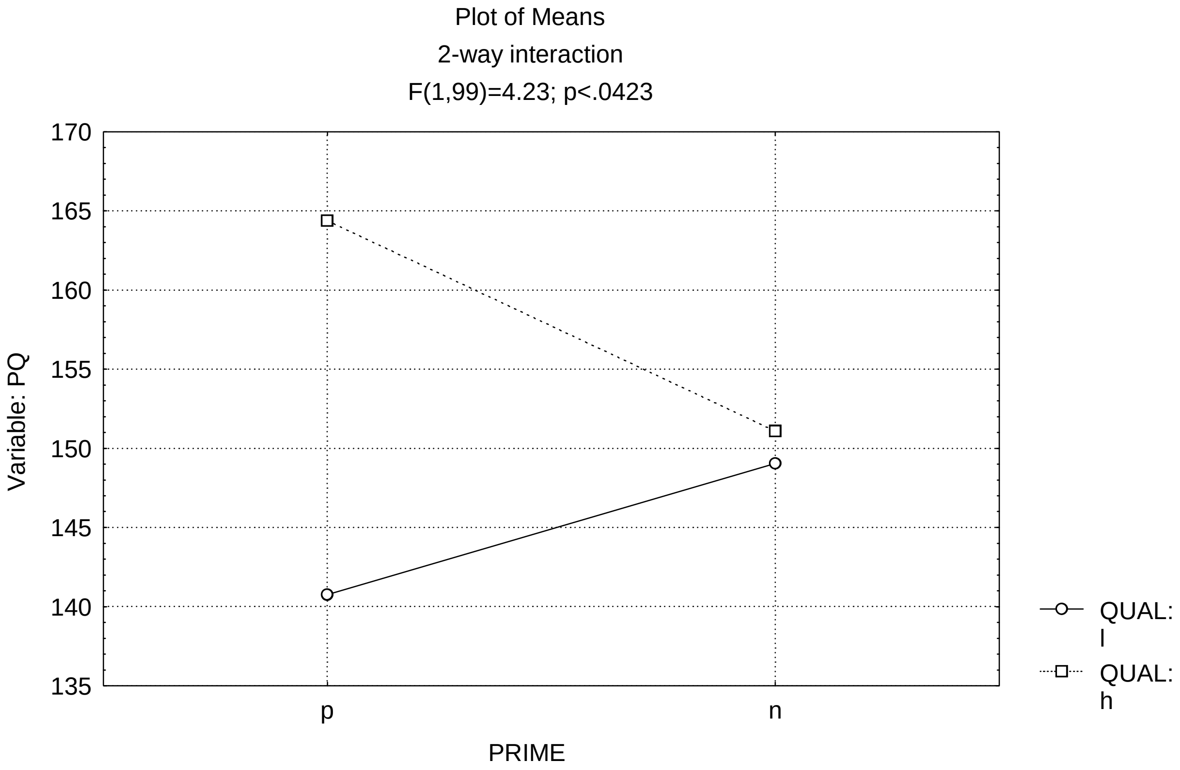


Figure 8-5: Means plot of the interaction between stimulus quality and priming on PQ

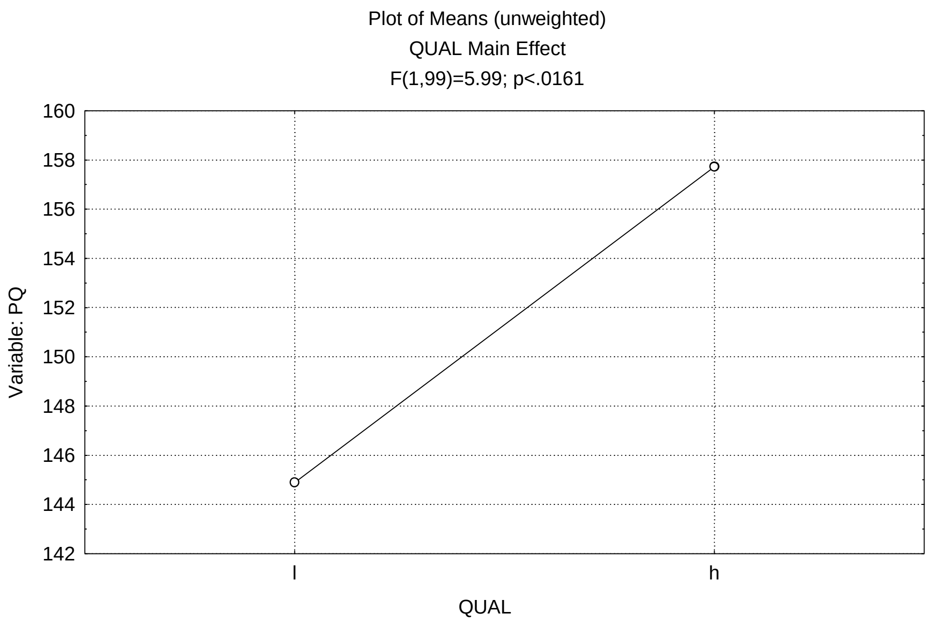


Figure 8-6 Means plot of the main effect of stimulus quality on PQ

8.7.5 Post-hoc analyses of the interaction between Stimulus Quality and Priming as independent variables and SUS and PQ as dependent variables

The interaction results reported in 8.7.4 above prompted a series of interesting post-hoc investigations. These included a focused look at the actual differences in means for each of the dependent variables at each of the priming and quality levels. Of the three dependent variables presented, COCI is the least interesting, having shown no interaction. However, PQ and SUS, apart from showing a remarkable similarity in their pattern of effects, each show an interaction effect between stimulus quality and priming. It is on these interactions on which this section focuses.

Interaction effect between Stimulus Quality and Priming on SUS: A series of protected t-tests was run to determine which means differences were statistically significant at each of the independent variable levels. The difference between Priming at the 'n' level (VE-irrelevant priming) and at the 'p' level (VE-relevant priming) shows significance for both the 'h' level of Stimulus Quality (high quality environment) and at the 'l' level (low quality environment). At the 'h' level, the t value is 2.18 (df = 49 $p < 0.04$) while at the 'l' level the t value is -2.34 (df = 50 $p < 0.03$). Furthermore, the difference between Stimulus Quality at the 'h' level (high quality environment) and at the 'l' level (low quality environment) shows a significant difference at the 'p' level of Priming (VE-relevant priming) with an exceptionally high t-value of 3.995 (df = 49, $p < 0.0003$); however, the difference between Stimulus Quality at the 'h' level and at the 'l' level is not significant at the 'n' level of priming (t = 0.07, df = 50, $p > 0.9$). Details of these analyses are presented in Table 8.7.

t-test conditions	Means tested	t value	df	*p*
Stim Qual = ‘h’	Priming ‘n’ vs. Priming ‘p’	**2.1832**	**49**	**0.0338**
Stim Qual = ‘l’	Priming ‘n’ vs. Priming ‘p’	**-2.3425**	**50**	**0.02318**
Priming = ‘n’	Stim Qual ‘l’ vs. Stim Qual ‘h’	0.0698	50	0.9446
Priming = ‘p’	Stim Qual ‘h’ vs. Stim Qual ‘l’	**3.9954**	**49**	**0.000217**

Table 8.7: post-hoc t-tests on SUS means at various independent variable levels. Significant tests ($p < 0.05$) in bold

Interaction effect between Stimulus Quality and Priming on PQ: As with the SUS analyses, protected t-tests were run to investigate the significance of means differences at each of the independent variable levels. In the case of the SUS, only 2 of the 4 tests are significant. At the ‘h’ level of Stimulus Quality (high quality environment), the difference between the PQ means at the ‘n’ level of Priming (VE-irrelevant priming) and at the ‘p’ level of priming (VE-relevant priming) is significant with t = 2.19 (df = 49, $p < 0.04$). The other significant difference is found at the ‘p’ level of Priming (VE-relevant priming) between the mean PQ scores at the ‘h’ level of Stimulus Quality (high quality environment) and the ‘l’ level of Stimulus Quality (low quality environment) having t value of 3.42 (df = 49, $p < 0.01$). The other two differences are not significant. The difference between the ‘n’ level of Priming and the ‘p’ level of priming at the ‘l’ level of Stimulus Quality was not significant (t = 0.97, df = 50, $p > 0.3$). The difference between the ‘l’ level of Stimulus Quality and the ‘h’ level of Stimulus Quality at the ‘n’ level of Priming is also not significant (t = 0.26, df = 50, $p > 0.75$). This information is summarized in Table 8.8.

t-test conditions	Means tested	t value	df	*p*
Stim Qual = ‘h’	Priming = ‘n’ vs. Priming = ‘p’	**2.1915**	**49**	**0.0332**
Stim Qual = ‘l’	Priming = ‘n’ vs. Priming = ‘p’	0.973	50	0.335
Priming = ‘n’	Stim Qual = ‘l’ vs. Stim Qual = ‘h’	0.26	50	0.79
Priming = ‘p’	Stim Qual = ‘h’ vs. Stim Qual = ‘l’	**3.42112**	**49**	**0.0013**

Table 8.8: post-hoc t-tests on PQ means at various independent variable levels. Significant tests ($p < 0.05$) in bold

8.7.6 Reliability analyses of scales used

To determine the internal consistency of the scales used, Cronbach’s alpha coefficient as well as item-total correlations were calculated for each of the scales. Cronbach’s alpha coefficient is a measure of the extent to which the items on a scale correlate with each other (Cronbach, 1960), and in the special case of a scale which measures only one construct, can be interpreted as estimating the degree to which the items agree on their estimate of the value of that variable (Anastasi, 1982). Thus, a high alpha value indicates a scale which is less prone to measuring noise. Item-total correlations represent the correlation between a particular item and the scale’s total excluding that item. Item-total correlations show the degree to which a particular item agrees with the scale as a whole. An item with a low item-total correlation suggests the measurement noise rather than of the construct that is being measured by the rest of the scale as a whole.

Reliability of the PQ: With 101 observations, the PQ shows an acceptably high Cronbach's alpha value of 0.9033 (standardized alpha value of 0.904). The item-total correlations are shown in Table 8.9. It should be noted that most of the correlations are in the 0.5 region. The average inter-item correlation is 0.235.

Item	Item-total correlation	Item	Item-total correlation	Item	Item-total correlation
PQ1	.494616	PQ12	.590483	PQ23	.364163
PQ2	.543138	PQ13	.328263	PQ24	.480119
PQ3	.620415	PQ14	.620148	PQ25	.644177
PQ4	.539336	PQ15	.559128	PQ26	.270850
PQ5	.670877	PQ16	.411539	PQ27	.223408
PQ6	.552051	PQ17	.198751	PQ28	.473694
PQ7	.567582	PQ18	.636567	PQ29	.302706
PQ8	.622474	PQ19	.202746	PQ30	.439059
PQ9	.436855	PQ20	.170757	PQ31	.239283
PQ10	.590865	PQ21	.422689	PQ32	.503193
PQ11	.563394	PQ22	.256610		

Table 8.9: Item-total correlations for PQ (n=101)

Reliability of the SUS: The SUS has only 6 items, so we can expect lower internal consistency (Anastasi, 1982). This is confirmed by a low Cronbach's alpha value ($\propto$ = 0.7695, std $\propto$ = 0.77, n=101). The average inter-item correlation is 0.3775. Table 8.10 shows the item-total correlations.

Item	Item-total correlation
SUS1	.700706
SUS2	.736349
SUS3	.487607
SUS4	.271800
SUS5	.400550
SUS6	.527470

Table 8.10: Item-total correlations for SUS (n=101)

Reliability of the COCI: The COCI has few items, and in line with measurement theory, it shows low internal consistency (Anastasi, 1982). It is not appropriate to calculate Cronbach's alpha for the COCI, as it is not a continuous scale, but rather a dichotomous one (selected themed word/did not select themed word). We thus calculate the Kuder-Richardson 20 (KR20) value, which is equivalent to the Cronbach's alpha for dichotomous data. The KR20 value is 0.74 (n=128). The average inter-item correlation is 0.224. Table 8.11 shows the item-total correlations.

Item	Item-total correlation
COC1	.317278
COC2	.516905
COC3	.331578
COC4	.304008
COC5	.309187
COC6	.344076
COC7	.460788
COC8	.474521
COC9	.297962
COC10	.429375

Table 8.11: Item-total correlations for COCI (n=128)

<u>*Reliability of the ITQ:*</u> According to its authors (Witmer & Singer, 1990), the ITQ has more than one factor, so it should not show a high Cronbach alpha value. However, the Cronbach alpha value is 0.8 (std $\propto$ = 0.801, n=44), which is higher than one would expect from a multi-factorial scale. The average inter-item correlation is 0.11. The item-total correlations are show in Table 8.12.

Item	Item-total correlation	Item	Item-total correlation	Item	Item-total correlation
ITQ1	.323684	ITQ13	.335411	ITQ25	.289553
ITQ2	.501146	ITQ14	.252606	ITQ26	.213596
ITQ3	.086090	ITQ15	.458851	ITQ27	.274783
ITQ4	.536089	ITQ16	.550660	ITQ28	.149905
ITQ5	.189501	ITQ17	.209872	ITQ29	.456395
ITQ6	.482474	ITQ18	.096614	ITQ30	.369215
ITQ8	.430204	ITQ19	.418629	ITQ31	.448047
ITQ9	.496124	ITQ21	.340194	ITQ32	.576710
ITQ10	.241416	ITQ22	.211616	ITQ33	.213210
ITQ11	.159359	ITQ23	.280761	ITQ34	.019286
ITQ12	-.049047	ITQ24	.049369		

Table 8.12: Item-total correlations for ITQ (n=44)

8.8 Discussion of results

The data collected in this experiment provide a large amount of information, which relates not only to the hypotheses outlined in 8.1, but also about the psychometric properties of the PQ, ITQ and SUS. It also provides some validation information with regard to the COCI. This section will briefly discuss the results outlined in 8.7, and relate them to the hypotheses outlined in 8.1.

8.8.1 Evidence for hypothesis 1: stimulus quality affects presence

The results show evidence is support of this hypothesis. Both the PQ and SUS show a significance increase in the high stimulus quality condition compared to the low stimulus quality condition. This effect is only evident in the priming relevant condition, so a certain amount of reserve is required when concluding. The COCI shows no effect with regard to stimulus quality, which might suggest a certain amount of doubt to any conclusion, but this is tempered by the fact that the PQ and SUS, both established measures of presence, are in agreement while it is the COCI, which is an unvalidated measure, that is in dissent. The parsimonious conclusion is therefore to disregard the COCI finding and conclude that presence is affected by stimulus quality. It seems that the hypothesis that stimulus quality affects presence is only true under certain conditions, one of which is the user's state of mind at the time of VR exposure. This point is dealt with in more detail in 8.8.3 below.

8.8.2 Evidence for hypothesis 2: priming affects presence

There seems to be little evidence for this hypothesis. The PQ and SUS are not affected by priming directly. The COCI, however, reacts strongly to the priming manipulation. COCI scores are significantly higher in the priming relevant condition as compared to the priming irrelevant condition. As with the effect described in 8.8.1 above, it would not be correct to conclude that presence is affected by priming, as the two established presence measures do not provide evidence to this effect. The fact that COCI is affected by priming should not be given too much credence, as this scale as yet unvalidated. Priming does have an effect on presence, but that effect is subtle, as discussed in 8.8.3 below.

8.8.3 Evidence for hypothesis 3: the interaction of priming and stimulus quality affects presence

The PQ and SUS measures provide important support for this hypothesis, although the COCI does not (the statements made above in 8.8.1 and 8.8.2 about the validity of the COCI apply here also). The PQ and SUS show a strong interaction, with presence scores at both levels of stimulus quality at the 'relevant priming' level differing significantly from each other, but those at the 'irrelevant priming' level not differing from each other. Furthermore, priming seems to have a beneficial effect in the high stimulus quality condition, and a detrimental effect in the low stimulus quality condition. This occurs in the SUS, although in the PQ the effect is slightly different – the presence mean in the high stimulus quality is improved by priming, but priming has no effect in the low stimulus quality condition. This complex behaviour suggests that presence is not simply a product of the sum of all the information available about the environment. Priming seems to act as a mediator variable rather than a predictor itself; it does not change presence scores directly (as shown in Section 8.8.2), but rather acts as a catalyst for other variables such as stimulus quality.

8.8.4 Evidence for the validity of the Contents of Consciousness Inventory

The strategy for validating a new scale is based on two broad strategies (Gregory, 1991):

(a) comparing the behaviour of a new scale to the behaviour of a scale of the same construct which is known to be valid

The COCI does not behave in a way that is consistent with the SUS and PQ, which represent the currently accepted practice of presence measurement. This is evidenced by the extremely low correlations between COCI and either SUS or PQ. In contrast, the PQ and SUS correlate extremely well with each other, suggesting that they are measuring comparable constructs.

(b) determining if the new scale behaves in a way consistent with the theory of the construct

According to our current understanding of presence, better display systems should lead to higher presence scores. The PQ and SUS both show main effects on stimulus quality, confirming this idea. The COCI however, does not show a main or interaction effect on stimulus quality. A second line of evidence that can be investigated in this regard, is the relationship between immersive tendencies and presence, proposed by Witmer & Singer (1998). The COCI shows no convincing relationship with the ITQ, which further reduces its validity. However, neither the PQ nor the SUS show any relationship with the ITQ either, leading to some doubt as to the utility of the ITQ/presence relationship as a method for validating presence measures.

From the evidence presented above, it seems clear that the COCI is not a valid measure of presence. The COCI detects the priming manipulation quite well, and is not affected by stimulus quality manipulations, so it would be more appropriate to refer to the COCI as a measure of priming rather than a measure of presence.

8.8.5 Reliability of the presence measures

There is no pre-scribed way to interpret a Cronbach's alpha value (Gregory, 1991), apart from "higher is better". Various authors suggest minimum acceptable thresholds, although the values they suggest are always quite high; for instance, Gregory (1991) suggests a minimum alpha of 0.9, and Anastasi suggests 0.8. Similar constraints apply to the interpretation of the Kuder-Richardson 20 formula. Based on these criteria, only the PQ performs as expected. The SUS and COCI fall short of the 0.8 mark, while the ITQ gives a borderline performance. Having a low level of internal consistency does not render a test completely useless. Rather, it implies that the test is more prone to measuring noise. The values derived from a test with low internal consistency will be less accurate, and the confidence of the statistical inferences will be reduced. With the results provided by these analyses, it seems that the PQ provides the most accurate presence readings. The measurements taken using the SUS, COCI and ITQ should thus be regarded with due skepticism.

Chapter 9

Experiment 2

9.1 VE presentation method and priming: Hypotheses

The purpose of this experiment is to test the idea that presence comes about not as a consequence of perception, but rather due to higher level mental processes. This will be done by comparing presence scores from participants immersed in graphics-based VEs to those immersed in text-based VEs. The specific hypotheses which are going to be tested are:

(a) Text-based VEs can produce presence as effectively as graphics-based VEs
(b) The effect of (a) above will be magnified under conditions of VE-relevant priming

The first hypothesis is derived from the notion that text-based VEs can be regarded as another level of stimulus quality (as defined in experiment 1). Provided that the text descriptions provide enough detail, a text based display could be considered as lying between the 'h' and 'l' graphics based levels. Graphics-based VEs represent "raw data" with regard to perception (the scene represented is simply what a virtual observer would see standing in a particular spot in the VE), while text-based VEs, which consist of written descriptions of a space, represent "pre-processed" data, which does not include all the raw information about the environment available in a graphics-based VE, but rather only those aspects which the author of the text descriptions thought relevant or important. If both of these cases produce comparable levels of presence, then this finding will support the notion that presence is not the result of direct perception only. The second hypothesis is derived from the results of experiment 1, which suggest that differences in presence between various conditions is magnified under conditions of VE-relevant priming. If text-based VEs also display this relationship to priming, then it will reinforce the idea that presence comes about as a product of cognitive processes which function at a level more abstract than perception, as the method of VE presentation would only have a small effect on the subsequent experience by the user.

In this experiment, the graphics-based VEs were the two monastery VEs used in experiment 1. The text-based VE is a implementation of the same VE used in the graphics-based system, but implemented in the TIVE text VE display system (cf. Appendix B for a description of this system). The descriptions of the rooms in the text VE version were built by an expert user after exploring the graphics version of the monastery VE. When a still image of a room was used in the text-based display system, this image was created by reducing the quality of a screenshot of the graphics based system. The text based system also allowed the user to perform the same number of actions as the graphics based system, with the exception of looking around in a room (for obvious reasons). By creating this correspondence between the text-based VE and the graphics-based VE, it becomes possible to make direct comparisons in presence levels between the text-based system and graphics-based system, as these can be regarded as different levels of the same variable, namely stimulus quality.

9.2 Variables

The same variables as used in experiment 1 (described in Chapter 8) were used. The main variables in this experiment are *presence*, *stimulus quality* and *priming*. Secondary variables used are *immersive tendencies* and *fun experienced in the virtual environment*.

9.2.1 Definition of variables

The variables used have the same definitions as those used in Chapter 8. These are:

- *Presence:* this is a complex concept, and is discussed in depth in Chapter two. Presence is used as a dependent variable in this experiment.

- *Stimulus quality:* this is the quality with which the virtual environment is rendered. It implies how closely the virtual rendering matches the virtual environment. This can be understood in several ways. One way of understanding this concept is as photo-realism. The higher the stimulus quality, the more photo realistic the rendering, although stimulus quality also includes the notion of multi modality ("photo-realism" for sound, etc as well as for images). Another way of understanding stimulus quality is through the notion of information load. A rendering of an environment which carries more environmentally relevant information has higher stimulus quality that one which carries less environmentally relevant information.

- *Priming:* this is the notion of providing the participant with information which is relevant to the virtual environment before they experience the environment. The idea of priming has existed in cognitive psychology in the constructive perception school as a way of influencing identification of ambiguous stimuli for many years, and is quite well understood.

- *Immersive tendencies :* this variable is defined by Witmer and Singer (1990). This is the capacity of a person to become immersed, be it in VR, cinema or any other immersive medium. Being a property of a person rather than a system, it is independent of any particular display device or medium, but some studies show that it is a predictor of presence. Its measurement is based on the participant's previous experiences with immersive media.

- *Fun experienced in the virtual environment:* this is the enjoyment experienced by the participant during the time of their immersion in the virtual environment. It is based on the participant's memory of their experience in the virtual environment.

9.2.2 Operationalization of variables.

- *Presence:* three measures of presence were used, namely the Presence Questionnaire of Witmer & Singer, the questionnaire of Slater, Usoh & Steed and the Contents of Consciousness Inventory introduced in this document. Presence is used as a dependent variable.

  The Presence Questionnaire (PQ) consists of 32 semantic differential items, each giving a score in the range from 1 to 7. In 28 of the items, a score of 1 indicates a response consistent with low presence, while a score of 7 indicates a score consistent with high presence. The remaining four PQ items (items 19, 22, 23 and 26) are reversed, with a score of 1 consistent with high presence, and a score of 7 being consistent with low presence. For the sake of simplicity, after administering the PQ, we reversed the response of these four items so that a higher total PQ score can always be understood as indicating more presence. A participant's presence score on the PQ is simply the sum of their responses on the 32 items. A completed PQ yields a score from 32 to 224. The PQ is reproduced in section G.2 of Appendix G.

The Slater, Usoh & Steed questionnaire (SUS) consists of 6 semantic differential items, each giving a score in the range from 1 to 7. For each item, a score of 1 indicates a response consistent with low presence, while a score of 7 indicates a response consistent with high presence. A participant's presence score on the SUS is simply the sum of their responses on the 6 items. A complete SUS produces a score of 6 to 42. The SUS is reproduced in section G.1 of Appendix G.

The Contents of Consciousness Inventory (COCI) consists of 10 forced choice items, each giving a score of 1 or 0. A score of 1 indicates that the participant has selected the option which matches the theme of the virtual environment, and 0 indicates that the participant has not selected the option that matches the theme of the virtual environment. The COCI score is simply the sum of all COCI items. A complete COCI gives a score ranging from 0 to 10. The COCI items used are listed in sections F.2 and F.3 of Appendix F.

- *Stimulus Quality:* In this experiment, stimulus quality is an independent variable, manipulated into three levels: 'l' or low stimulus quality, 'h' or high stimulus quality and 't' or text implementation. In all the levels, the same virtual environment and geometry is presented. In the low stimulus quality level (the 'l' level), the environment is not textured, Lambertian lighting is used, and no sound is present. In the high stimulus (the 'h' level), radiosity is used and the geometry is textured. There is also sound, both environmental (sounds which emanate from no particular place and serve as an aural backdrop, such as the sound of thunder) and positional sound (sounds originating from a particular place, such as the sound of ringing from a telephone). In the text implementation, the system displayed the rooms as text descriptions of about 150 words, with some rooms (70% of them) including a still image of the room. These still images were 200x180 pixels in size, and displayed in 8 bit colour. The text implementation included no sound.

- *Priming:* priming is an independent variable in this experiment and is therefore manipulated rather than measured. Two levels of priming exist: a VE relevant priming condition (abbreviated 'p') and a VE irrelevant priming condition (abbreviated 'n'). In the VE relevant condition ('p' level), the participant is allowed to examine, prior to the VR experience, a booklet which is related in theme to the virtual environment they are about to experience. In the VE irrelevant priming condition ('n' level), the same manipulation occurs, but the booklet is not related to the virtual environment they are about to experience. Details of the contents of the booklets are stated in section 9.4.6. The booklets are reproduced in sections H.1 and H.3 of Appendix H.

- *Immersive tendencies:* this variable is measured by means of the Immersive Tendencies Questionnaire (ITQ) compiled by Witmer & Singer (1990). The ITQ consists of 32 semantic differential items, each giving a score in the range 1 to 7, with a response of 1 being consistent with low immersive tendencies, and a score of 7 being consistent with an individual who tends to be highly immersive. A participant's ITQ score is simply the sum of all their responses. A complete ITQ provides a score which ranges between 32 and 224. The ITQ is reproduced in section G.3 of Appendix G.

- *Fun experienced in the environment:* this variable is measured with a single Likert-type item, in the range 1 to 10. A score of 1 is consistent with having felt no fun, while a score of 10 is consistent with having had a lot of fun. A participant’s score is simply their response to this item, in the range 1 to 10.

9.3 Design

This study made use of a simple one-way design, with 2 cells (primed and not-primed). The design did not make use of repeated measures.

9.4 Materials

9.4.1 Venue

The experiment was conducted in a dedicated room in which lighting and extraneous noise could be controlled by means of the door and window blinds. Inside this room, three computers were placed so that up to three participants could be run simultaneously. Partitions were placed between each computer so that participants could not see each other or their displays (although they entered the room together and were aware that others were in the room with them). The experimenter could stand behind a partition on one end of the room and observe the participants without them being distracted by his observation.

9.4.2 Computers

3 computers were used for this experiment, to allow 3 participants to be run simultaneously. All computers were equipped with the same hardware and software, namely

- AMD Athlon 700 MHz processor
- 128 MB RAM
- GeForce 2 MX 32MB graphics card
- 17” monitor displaying a 320x240x8 image (with user-event triggered refresh)
- Windows 2000
- Sound Blaster sound card
- Stereo headphones
- Keyboard and mouse
- For the graphical TIVE v1.0 (cf. Appendix B for a description)

9.4.3 VR display system

This experiment used our TIVE v1.0 tool to display the virtual environments. This tool is described in more detail in Appendix B.

9.4.4 Participants' task

The participants were given a simple task to perform during their exposure in the VE. Inside the VEs, a series of small boxes, called tokens, were placed by the experimenter in various locations. Only one of the tokens was visible at any one time. The participant's task was to search for these tokens and collect them by stepping into the room containing the token. When a token had been found, it disappeared, and the next token in the series would appear in another location. To help the participants find the tokens, a "token scanner" was available. This scanner showed a participant their position relative to the current token. By pressing the 'T' key, participants would bring up a text window which would give the bearing and relative height to the token in terms of the compass directions. There was no indication of distance to the token.

The purpose of this task was simply to keep the participant's attention focused on the VE, and to ensure that the participants did not stay in a small area of the VE. By placing a token in each major area of the VE, the experimenter could ensure that most participants visited most areas of the VE, without fear of them becoming lost or wandering in circles. No measurements were taken with regard to the task.

9.4.5 Virtual environments

Three TIVE virtual environments were used in this experiment:

- train.tve - This environment is abstract, and does not represent any particular theme. It is a text implementation of the train.cfg environment used in experiment one. It is a brick building with 12 sparsely furnished rooms on three levels, and is used to train the participants in the use of the TIVE tool. No tokens are defined. No COCI items are defined.

- traincoc.tve - This environment is used to teach the participants about the COCI and the token scanner. It is the same as train.tve except that 13 tokens and a set of 6 COCI items is defined.

- monastery.tve – The monastery. This environment represents a medieval monastery and chapel in the countryside. The environment features 27 locations on three levels, most of which (70%) are accompanied by a still image. 10 COCI items are defined. There are also 28 tokens defined.

  The monastery.tve environment is a text implementation of the mon_hi.cfg environment used in experiment one. This was done so that data collected in this experiment could be directly compared with data collected in the monastery environment in experiment one.

9.4.6 Priming materials

Two priming booklets were used, which were identical to those used in experiment 1 (described in Chapter 8); one related in theme to the monastery virtual environment, and a neutral booklet with a theme unrelated to the monastery virtual environment. Each booklet contained a piece of text, approximately 1000 words long, as well as approximately 4 pictures. In the case of the monastery themed booklet, the pictures were printed in colour (the neutral booklet's pictures were not printed in colour as a cost saving measure). The booklets have been reproduced as an appendix to this document.

- Monastery booklet - This booklet contained a brief history of the monastic movement in medieval times, and a brief synopsis of everyday life in a monastery. The text was assembled from several web pages on the subject, and consisted of 963 words, and five pictures. Two of the pictures were of monasteries in England and France, and the other three were scans of illuminated religious scripts from the medieval period. See section H.1 of Appendix H for a re-print of this booklet.

- Neutral/non-priming booklet – This booklet contained a description of a driving steam trains. It explains the procedures of starting the train, as well as some personal anecdotes by the authors, and was assembled from several web pages on the subject. The booklet consisted of 1051 words and 3 pictures. The pictures showed steam trains, as well as the controls of a locomotive. The booklet is included in section H.3 of Appendix H.

9.4.7 Scales & Questionnaires

Four separate sets of scales were administered to each participant. These included a COCI set (themed to the monastery), which was administered electronically during the exploration stage. After the exploration stage, a set of questionnaires printed on paper was administered. The questionnaire booklet, as presented to the participants, has been reproduced as an appendix to this document.

The questionnaire set included the following scales, in this order:

- Slater, Usoh & Steed (2000) scale
- Presence Questionnaire (Witmer & Singer, 1998)
- Form 100 (developed for this experiment)

9.4.8 Experimental schedule

Due to the simple design of this study, the schedule simply assigned participants randomly into the priming condition of the non-priming condition. The schedule was designed to assign the same number of observations in each condition.

9.5 Participants

Posters advertising the experiment were placed around the Computer Science building at the University of Cape Town. This building is used not only by computer science students, but also by undergraduates of the Science faculty, the Commerce faculty and the Health Sciences faculty. The posters asked for participation in a virtual reality experiment for a payment of 20 Rand. The posters did not express preference for any particular group. Volunteers were asked to write their name and contact telephone number on a sign-up sheet outside the experimental venue. The sign-up sheets contained details of the times available for participation.

Demographic details of the participants were not recorded; however, all were volunteers. Although details were not recorded, the group included both men and women (with a much higher proportion of men) of various ethnic groups. Almost all participants were in their early twenties.

A total of 48 volunteers wrote down name and contact telephone number on the sign up sheet. Of those, 25 actually arrived at the agreed time and participated in the experiment. The others did not appear at the experimental venue. This represents a 54.2% response rate.

9.6 Procedure

The procedure of this experiment is quite similar to that of experiment one, the main difference being that the participants explored only one virtual environment. This similarity was created so that data collected in this experiment could be directly compared with the data collected in experiment one.

9.6.1 Introduction stage

The participants were taken into the experimental venue, and told that they would be taking part in a study to evaluate a possible solution of displaying virtual environments in impoverished displays such as PDAs and cellphones. The experimenter explained that the system was text based only, as well as being non-realtime. The participants were told that their role in the evaluation of the system was to explore a virtual environment, taking in the sights along the way.

9.6.2 Training stage

The TIVE tool was started displaying the train.tve environment. The movement interface (excluding details on the tokens and the COCI) was explained, and the participants were allowed to practice this until the experimenter was satisfied that all of the participants understood the system (this usually took no longer than three minutes). The TIVE tool was shut down, and the participants were told of the tokens and token scanner. They were told that the tokens existed to help them explore the environment more thoroughly, and that collecting the tokens was not their main task; the main task is to carefully explore the environment. The COCI was then explained to the participants. The TIVE tool was then started with the traincoc.tve environment, which includes both tokens and COCI items. The participants were allowed to explore freely until the first COCI item appeared, at which point the experimenter again explained how to respond to the items. The participants were then allowed to practice exploring the environment and responding to COCI items under the supervision of the experimenter. When the experimenter was satisfied that all the participants understood the interface and COCI, the TIVE tool was shut down (this usually took no longer then 5 minutes).

9.6.3 Priming stage

The participants were told that the training was complete and that the experiment was about to begin. They were informed of the procedure, namely that they would be given a booklet to read, followed by exploring a virtual environment, followed by filling out a questionnaire. They were reminded that their primary task in the virtual environment was exploration, and not collecting of the tokens. They were then told that they would only have 5 minutes to study the booklet they were about to be given, but that they should not rush through it. They were told that reading slowly and carefully was more important than finishing the booklet. They were then given the booklet as dictated by the experimental schedule, and allowed to read for five minutes.

9.6.4 Exploration stage

The TIVE tool was started with the monastery.tve environment, and the participants were instructed to begin exploring. The experimenter remained in the room, but observed the participants from a hidden position so that the participants were not distracted. Once all of the participants had completed the entire set of COCI items (which took between 11 and 15 minutes), the exploration stage was concluded.

9.6.5 Questionnaire stage

The TIVE tool shut down, and the participants were handed the questionnaire set. The participants were then given time to complete the entire set of questions, which usually took between 10 and 15 minutes.

9.6.6 Completion and preparation stage

Once the participants had completed the questionnaires, they were asked to sign a receipt, and were paid 20 Rand for participating. When the participants had left, the experimenter checked the data output file in each computer for completeness, and prepared each computer for the next set of participants. The paper questionnaires were then coded and stored.

9.7 Analysis of results

In this section we present an analysis of the data collected in this experiment. We first present descriptive statistics for each of the variables, followed by an inferential analysis. In the section pertaining to inferential statistics, we have combined the data collected in experiment 2 (text-based VE data), with data from the monastery VE condition collected in experiment 1. Only the monastery VE data was used as the text-based VE only displayed the monastery text environment. As the conditions under which the data was collected in the two experiments were very similar, it is permissible to combine the data in this way. When the data are combined, the 'text' condition is regarded as another level of the s*timulus quality* variable. The analyses presented were conducted using Microsoft Excel 2000 and StatSoft Statistica 5.5.

9.7.1 Categorization of participants into conditions

This section includes the data collected in the monastery VE in experiment 1 as well as the data collected in experiment 2 (text condition). Each of the 78 participants contributed an observation to one of the cells. Table 9-1 shows the frequency of observations in each of the four cells of the design.

	Priming		
Stimulus Quality	Relevant priming ('p')	Irrelevant priming ('n')	*Row totals*
Low quality ('l')	13	13	*26*
High Quality ('h')	14	13	*27*
Text ('t')	14	11	*25*
Column totals	*41*	*37*	**78 (Grand Total)**

Table 9-1: Number of observations in each of the conditions

9.7.2 Descriptive statistics for COCI, SUS, FUN, PQ and ITQ

This section includes only the data collected in experiment 2 (the text condition). Each of the 25 observations was made on Witmer and Singer's Presence Questionnaire (PQ), Slater, Usoh & Steed's questionnaire (SUS) and the Contents of Consciousness Inventory (COCI). Furthermore, in the Form 100 questionnaire, participants were asked to rate their sense of fun on a scale of 1-10 (FUN). All of the 25 participants also completed Witmer and Singer's Immersive Tendencies Questionnaire (ITQ). The descriptive statistics for these variables are presented in Table 9-2.

Variable	Valid N	Mean	-95% CI	+95% CI	Min	Max	Std.Dev.
COCI	25	6.52	5.7559	7.2841	3	10	1.85113
SUS	25	22.96	21.0930	24.8270	14	32	4.5229
FUN	25	5.68	4.5962	6.7638	1	10	2.62552
PQ	25	119.28	107.8970	130.6630	64	181	27.57644
ITQ	25	157.16	148.8319	165.4881	129	195	20.17565

Table 9-2: Descriptive statistics for COCI, SUS, FUN, PQ and ITQ

9.7.3 Correlations between COCI, SUS, FUN and PQ

This section is based only on data collected in experiment 2 (text condition). To test if the various presence measurements are valid, we conducted a series of correlations to test the concurrent validity of these measures. Also, we looked for correlations between FUN and the various presence measures. A correlation matrix (n=25) was created from the variables FUN, PQ, ITQ, COCI and SUS. Table 9-3 presents a summary of the results, with significant correlations ($p < 0.05$) being in bold.

	COCI	SUS	FUN	PQ	ITQ
COCI	1.0	*-.12*	*-.03*	*.20*	*.19*
SUS	*-.12*	1.0	**.42**	**.56**	*.24*
FUN	*-.03*	**.42**	1.0	**.68**	*.33*
PQ	*.20*	**.56**	**.68**	1.0	**.58**
ITQ	*.19*	*.24*	*.33*	**.58**	1.0

Table 9-3: Correlations between COCI, SUS, FUN, PQ and ITQ.
Significant correlations ($p < 0.05$) in bold

9.7.4 Factorial ANOVA with Stimulus Quality and Priming as independent variables and SUS, PQ and COCI as dependent variables

This section includes the data collected in the monastery VE in experiment 1, as well as the data collected in experiment 2 (text condition). To test the effect of stimulus quality and priming on the measures of presence, a 2x2 factorial ANOVA was conducted for each of the three dependent variables (SUS, PQ and COCI).

COCI as the dependent variable – all effects: An investigation of the effects shows that there is no significant interaction between stimulus quality and priming ($F(2,72) = 0.33$ $p > 0.716$). Both main effects are significant; the main effect for stimulus quality ($F(2,72) = 7.71$ $p < 0.001$), as well as the main effect for priming ($F(1,72) = 4.511$ $p < 0.04$). A summary of these statistics is presented in Table 9-4. Means plots of the significant main effects are presented in Figure 9-1 and Figure 9-2.

	df Effect	MS Effect	Df Error	MS Error	F	p-level
Stim Qual	**2**	**43.25228**	**72**	**5.607968**	**7.712646**	**.000923**
Priming	**1**	**25.29770**	**72**	**5.607968**	**4.511027**	**.037113**
Interaction	2	1.87543	72	5.607968	.334423	.716857

Table 9-4: Summary table of effects for a 2x2 factorial ANOVA with COCI as the D.V. Significant effects ($p < 0.05$) marked in bold.

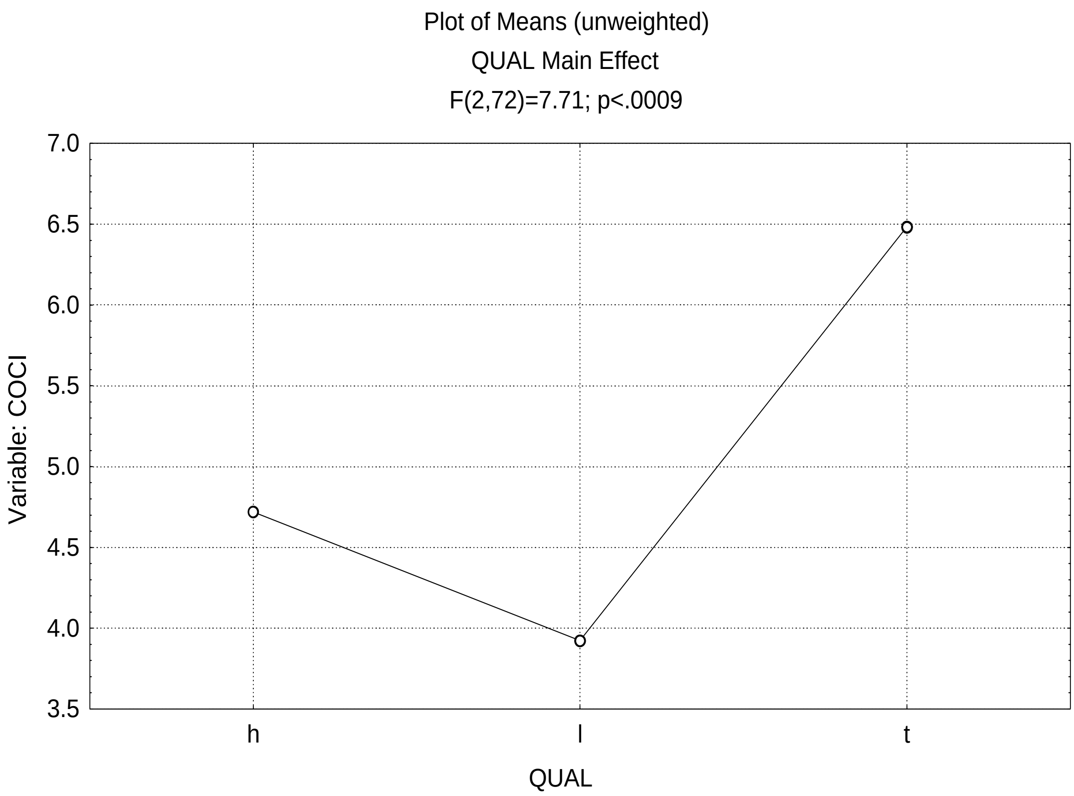


Figure 9-1: Means plot of the main effect of stimulus quality on COCI

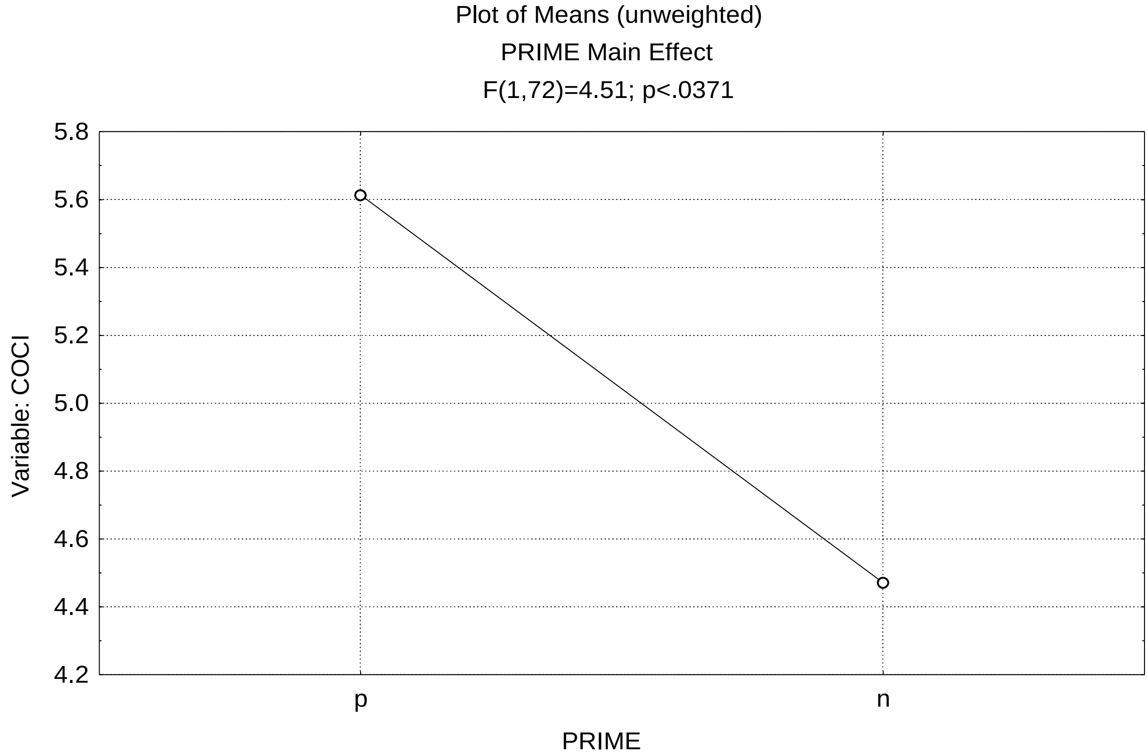


Figure 9-2: Means plot of the main effect of priming on COCI

<u>*SUS as the dependent variable – all effects:*</u> This variable presents a significant interaction between stimulus quality and priming (F(2, 72) = 5.87 $p < 0.005$). The means plot of this effect is shown in Figure 9-4. There is also a significant main effect on stimulus quality (F(2,72) = 4.58 $p < 0.02$). The means plot of this effect is shown in Figure 9-3. There is, however, no significant main effect in priming (F(1,72) = 2.74 $p > 0.1$). This effect information is summarized in Table 9-5.

	Df Effect	MS Effect	df Error	MS Error	F	p-level
Stim Qual	**2**	**212.5133**	**72**	**46.31329**	**4.588603**	**.013315**
Priming	1	127.3449	72	46.31329	2.749641	.101627
Interaction	**2**	**272.2772**	**72**	**46.31329**	**5.879030**	**.004315**

Table 9-5: Summary table of effects for a 2x2 factorial ANOVA with SUS as the D.V. Significant effects ($p < 0.05$) marked in bold.

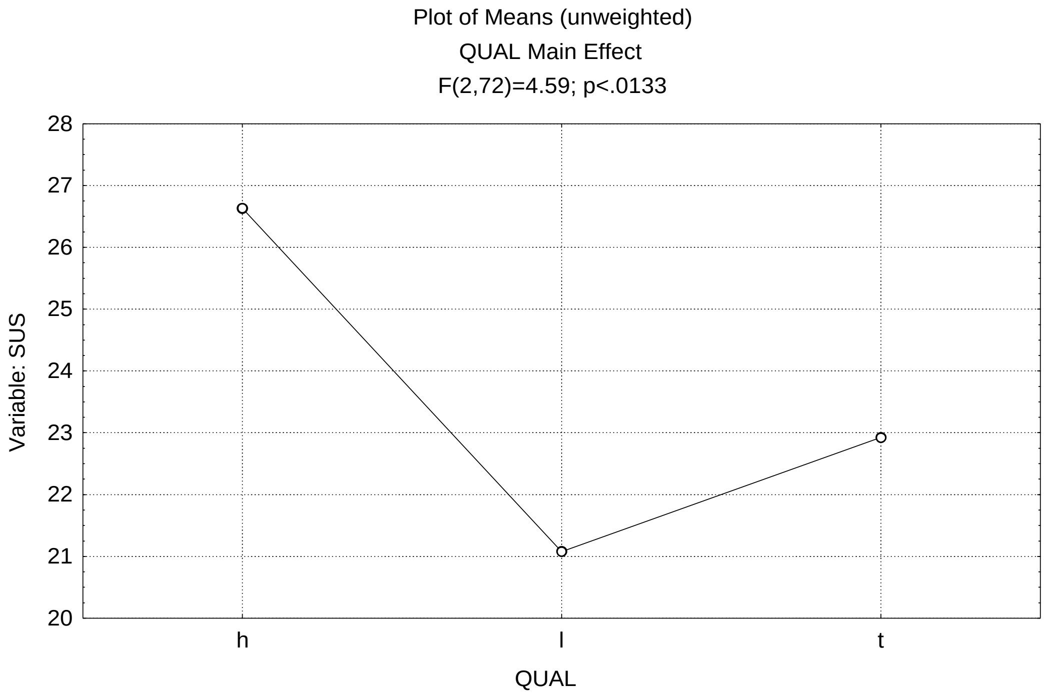


Figure 9-3: Means plot of the main effect of stimulus quality on SUS

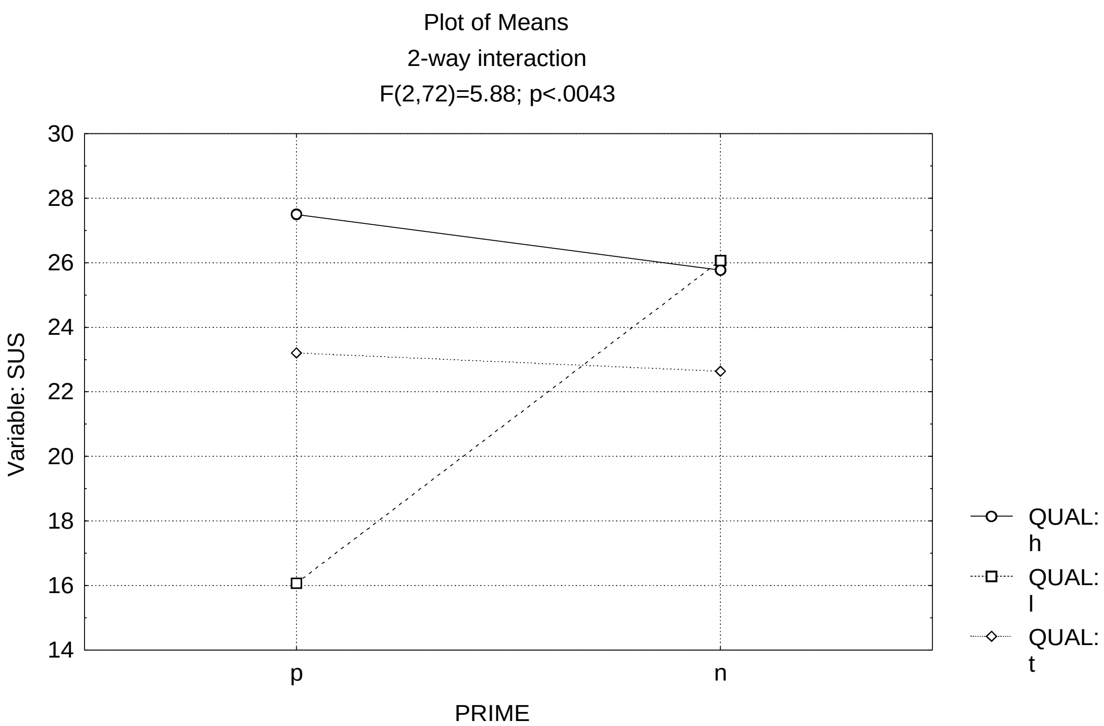


Figure 9-4: Means plot of the interaction between stimulus quality and priming on SUS

PQ as the dependent variable – all effects: This variable shows very few significant effects. The interaction between stimulus quality and priming is not significant ($F(2,72) = 0.813$ $p > 0.44$), neither is the main effect of priming ($F(1,72) = 0.536$ $p > 0.46$). The main effect of stimulus quality on PQ, however, is significant ($F(2,72) = 10.467$ $p < 0.001$). A means plot for this effect is presented in Figure 9-5. Table 9-6 presents a summary of the effects.

	df Effect	MS Effect	Df Error	MS Error	F	p-level
Stim Qual	**2**	**8141.906**	**72**	**777.8716**	**10.46690**	**.000102**
Priming	1	417.183	72	777.8716	.53631	.466342
Interaction	2	632.846	72	777.8716	.81356	.447310

Table 9-6: Summary table of effects for a 2x2 factorial ANOVA with PQ as the D.V. Significant effects (p < 0.05) marked in bold.

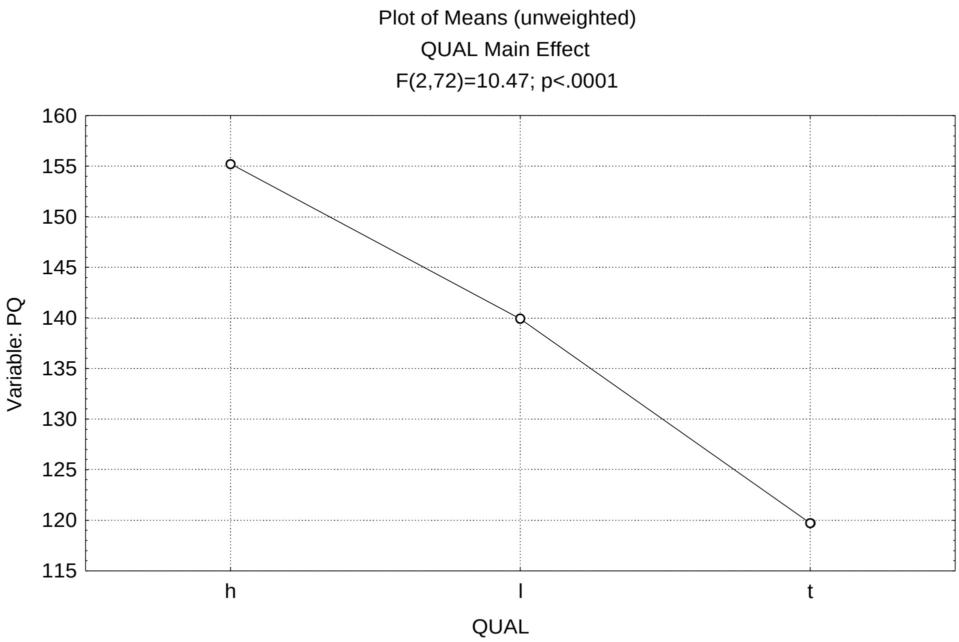


Figure 9-5: Means plot of the main effect of stimulus quality on PQ

9.7.5 Post-hoc analyses of the main effects if stimulus quality

In order to address the hypotheses proposes in 9.1, it is necessary to examine the difference in presence between the various levels of the stimulus quality factor. The ANOVAs show that the stimulus quality main effect is significant for all three presence scales, namely COCI, SUS and PQ. These post-hoc analyses were conducted by means of a series of protected t-tests.

Stimulus Quality main effect on SUS: These tests show that the 'h' level of stimulus quality is different from both 'l' ($t = 2.4501$, df = 51, $p < 0.02$) and 't' ($t = 2.2658$, df = 50, $p < 0.03$). However, the 'l' and 't' conditions are not significantly different ($t = 0.8956$, df = 49, $p > 0.37$). These results are summarized in Table 9-7, and can be appreciated graphically in Figure 9-5.

Means tested	t	df	*p*
Stim Qual = ‘h’ vs. Stim Qual = ‘l’	**2.4501**	**51**	**0.0177**
Stim Qual = ‘h’ vs. Stim Qual = ‘t’	**2.2658**	**50**	**0.0278**
Stim Qual = ‘t’ vs. Stim Qual = ‘l’	0.8956	49	0.3748

Table 9-7: Post-hoc tests of stimulus quality conditions on SUS. Significant results ($p < 0.05$) in bold

Stimulus Quality main effect on COCI: The pattern in the COCI is quite different from that in the SUS, as can be seen in Figure 9-1. Here, the ‘t’ level gives the highest mean score, being significantly different from both the ‘h’ level (t = 2.8587, df = 50, $p < 0.01$) and the ‘l’ level (t = 3.9935, df = 50, $p < 0.001$). The ‘h’ and ‘l’ level, however, produce equivalent scores (t = 1.13427, df = 51. $p > 0.25$). Table 9-8 shows these results.

Means tested	t value	df	*p*
Stim Qual = ‘h’ vs. Stim Qual = ‘l’	1.13427	51	0.2619
Stim Qual = ‘h’ vs. Stim Qual = ‘t’	**2.8587**	**50**	**0.00618**
Stim Qual = ‘t’ vs. Stim Qual = ‘l’	**3.9935**	**49**	**0.000218**

Table 9-8: Post-hoc tests of stimulus quality conditions on COCI. Significant results ($p < 0.05$) in bold

Stimulus Quality main effect on PQ: The PQ shows a distinct progression in presence scores, as seen in Figure 9-5, starting with the ‘h’ level, and moving down to ‘l’ and then ‘t’. The ‘h’ level is significantly different from both the ‘l’ level (t = 1.0175, df = 51, $p < 0.05$) and the ‘t’ level (t = 5.2002, df = 50, $p < 0.0001$). The ‘l’ level is also significantly different from the ‘t’ level (t = 2.4396, df = 49, $p < 0.02$). These results are summarized in Table 9-9.

Means tested	t value	df	*p*
Stim Qual = ‘h’ vs. Stim Qual = ‘l’	**2.0175**	**51**	**0.0489**
Stim Qual = ‘h’ vs.Stim Qual = ‘t’	**5.2002**	**50**	**0.000004**
Stim Qual = ‘t’ vs. Stim Qual = ‘l’	**2.4396**	**49**	**0.0183**

Table 9-9: Post-hoc tests of stimulus quality conditions on PQ. Significant results ($p < 0.05$) in bold

9.7.6 Post-hoc analyses of the interaction between Stimulus Quality and Priming as independent variables and SUS as the dependent variable

The results from this experiment show interaction only on SUS, unlike the results presented in experiment 1, which showed interactions on SUS as well as PQ. This section presents a series of protected t-tests which was done to investigate the specific effect of the interaction.

Interaction effect between Stimulus Quality and Priming on SUS: A series of protected t-tests was run to determine which means differences were statistically significant at each of the independent variable levels. Only 3 of the 9 tests are significant. The difference between the 'n' and 'p' levels of priming at the 'l' level of stimulus quality is significant (t = 3.1033, df = 24, $p < 0.004$). The difference between the 'h' and 'l' levels of stimulus quality at the 'p' level of priming is also significant (t = 3.5242, df = 25, $p < 0.0017$) as is the difference between the 'l' and 't' levels of stimulus quality at the 'p' level of priming (t = 2.4208, df = 25, $p < 0.025$). The details of these tests and the remaining, non-significant tests are presented in Table 9-10. These tests suggest a pattern similar to the interaction effect produced by SUS in experiment 1. In that case, the 'l' level of stimulus quality showed moderate scores at the 'n' level of priming, and lower scores at the 'p' level of priming, while the 'h' level of stimulus quality showed the opposite effect (moderate scores on 'n' and high scores on 'p'). In this experiment, the 'h' and 't' levels of stimulus quality show the same scores across priming conditions, while the 'l' level shows the same pattern as in experiment 1 – moderate scores at the 'n' level of priming and low scores at the 'p' level of priming.

t-test conditions	Means tested	t value	df	*P*
Stim Qual = 'h'	Priming = 'n' vs. Priming = 'p'	0.6417	25	0.5268
Stim Qual = 'l'	Priming = 'n' vs. Priming = 'p'	**3.1033**	**24**	**0.004**
Stim Qual = 't'	Priming = 'n' vs. Priming = 'p'	0.3111	23	0.7585
Priming = 'n'	Stim Qual = 'h' vs. Stim Qual = 'l'	0.1172	24	0.90767
Priming = 'n'	Stim Qual = 'h' vs. Stim Qual = 't'	1.3548	22	0.1892
Priming = 'n'	Stim Qual = 'l' vs. Stim Qual = 't'	1.5528	22	0.1347
Priming = 'p'	Stim Qual = 'h' vs. Stim Qual = 'l'	**3.5242**	**25**	**0.00166**
Priming = 'p'	Stim Qual = 'h' vs. Stim Qual = 't'	1.8158	26	0.0809
Priming = 'p'	Stim Qual = 'l' vs. Stim Qual = 't'	**2.4208**	**25**	**0.023**

Table 9-10: post-hoc t-tests on SUS means at various independent variable levels. Significant tests ($p < 0.05$) in bold

9.8 Discussion of results

This section briefly discusses the results presented in 9.7, and relates these to the hypotheses laid out in 9.1. The data collected in this experiment presents a complex picture, with each of the presence scales performing differently. This makes the results collected here less clear than hoped, as having no agreement between scales raises the possibility that the finding is not an effect of presence itself but rather of the scale used to measure it.

No attempt was made to analyze the data from this study to assess the reliability or validity of scales, as the sample size was not large enough for this purpose.

9.8.1 Evidence for hypothesis 1: Text-based VEs produce the same levels of presence as graphics-based VEs

The results show some support for this hypothesis, although each of the presence scales shows a different picture. If one ignores priming as a factor, the SUS shows that the text condition produces as much presence as the low quality graphics condition, while the COCI shows that the text condition produces the highest levels of presence. Finally, the PQ opposes the hypothesis by suggesting that the text condition produces the lowest of all the conditions. Faced with these results, it would not seem incorrect to conclude that very little support exists for this hypothesis.

However, if priming is brought into the analysis, then the findings become more promising. Unlike experiment 1, where both the PQ and SUS showed a priming/stimulus quality interaction, the results from this experiment show an interaction only for SUS, which makes the argument somewhat weaker. Nonetheless, when VE-relevant priming has occurred, participants in the text condition report presence levels which are not significantly different from the high quality graphics condition, and significantly better than the low quality graphics condition. This implies that under conditions of priming, the text condition is as effective as the high quality graphics condition. When the priming was VE-irrelevant, the three conditions (text, high quality graphics and low quality graphics) perform equally, in a pattern similar to that found in experiment 1.

The odd results produced by the COCI, which place text-based VEs significantly higher than graphics-based VEs in terms of presence production, might be due to an artifact of the scale itself. The COCI, as used in this experiment, was based on written language; the participants selected the word they thought stood out from a list of 4 words. The text-based VE is also based on written language, and this may lead to a confound with memory effects, as some of the words from the COCI item lists had appeared previously in the descriptions of the rooms. For example, the words “candle” and “wood” existed in room descriptions, and so may have been selected on this basis rather than due to any effect of presence. Sadly, there is not enough data to explore this possibility further without further experimentation.

One can summarize the discussion above by concluding that the hypothesis that text based VEs produce the same levels of presence as graphics based VEs does have sufficient support. With the SUS, text based VEs perform similarly to high-quality VEs, and the COCI text based VEs outperform graphics-based VEs in terms of presence production. A slight doubt as to the veridity of this hypothesis is raised by the PQ, which ranks the text based VEs lowest in terms of presence production.

9.8.2 Evidence for hypothesis 2: Priming magnifies the effect of presence in text-based VEs

This hypothesis aimed to show that text-based VEs would behave as the graphics-based VEs did in experiment 1, where priming was found to affect the presence scores by way of an interaction with stimulus quality. This experiment does provide much evidence to support this notion. In the COCI, the presence scores are affected directly by priming; that is, all three conditions benefited from priming, replicating the pattern observed in experiment 1. The text-based VE condition SUS scores show a true interaction, although there is no significant difference in SUS scores between VE-relevant and VE-irrelevant priming. However, this pattern is also exhibited by the high quality graphics condition, which shows no significant increase from the VE-irrelevant condition to the VE-relevant condition. The text condition thus shows the same behaviour across priming conditions as the high quality graphics condition. This behaviour itself quite different from that shown by the low quality graphics condition, which shows a significant drop from the VE irrelevant condition to the VE relevant condition.

This analysis seems to suggest that with respect to priming, text-based VEs behave in a way similar to high-quality graphics VEs, but quite differently to low quality graphics VEs. This finding does not satisfy the hypothesis exactly, as there is no significant improvement in the VE-relevant condition over the VE-irrelevant condition. However, this finding does suggest that text-based VEs behave in a way which is not significantly dissimilar from some graphics-based VEs, which in itself suggests that the mental processing of these two different modes of presenting VEs leads to similar products.

9.8.3 Caveat – Suitability of the PQ for text-based VEs

From the data collected in this study, it seems that the PQ might not be the most appropriate scale for use with text-based systems. The stimulus quality main effect on PQ showed the text condition to be significantly lower than both graphical conditions, but this result was not corroborated by either of the other two presence scales used. This behaviour can be attributed to the nature of the content of the items in the PQ. An investigation into the items reveals that most items ask directly about properties of the display system, for example, items 10 and 12, reproduced below:

> *10. How completely were you able to actively survey or search the environment using vision?*
>
> *12. How well could you localize sounds?*

This type of item refers less to the participant’s experience than it does to the specifications of the display system. In a text based system, where there are almost no visuals and no sound, all subjects are compelled to respond negatively to those items, even if the text based VE was having a profound impact on them. This type of item, which asks about the *purported causes of presence* rather than about the *sense of presence experienced* is not useful when investigating other immersive media which are not similar to those expected by Witmer & Singer, who implicitly impose a requirement that a medium be multimodal, interactive and realtime in order to produce presence. This same line of argument is taken by Slater (1999), who argues that Witmer & Singer’s scale to strictly measures the operationalization of presence rather than the variable itself, In cases such as text based VEs, the SUS, which asks directly about the experience rather than about the display system, may be more appropriate measure of presence.

Chapter 10

Evaluating the evidence for the connectionist model of presence

Chapter 7 outlined a strategy for evaluating the correctness of the connectionist model of presence, and this strategy was implemented by the empirical studies described in Chapter 8 and Chapter 9. This chapter reconsiders the evidence presented in these experiments, and evaluates the model in terms of the specific predictions made by the model in Chapter 7. This evaluation is then used as the basis for a critical evaluation and discussion of the model.

10.1 Empirical evidence for the interaction between conceptual layer activation and perceptual analyzer activation

Section 7.3.1 in chapter 7 outlined the specific results predicted by the model for the four novel conditions created by manipulating two levels of conceptual layer activation and two levels of perceptual analyzer activation. These four conditions were implemented in experiment 1 (described in chapter 8).

10.1.1 Prediction 1 – A difference between the *High stimulus quality* and *Low stimulus quality* conditions

The evidence for this prediction can be found in section 8.7.4 in Chapter 8. The analysis of the data showed a significant difference in this regard with the SUS and PQ scales, but not the experimental COCI. In the case of the SUS (Table 8.5) and PQ (Table 8.6), subjects in the *High stimulus quality* condition experienced more presence than in the *Low stimulus quality* condition. As two of the three scales used to measure presence in this experiment produced the predicted results, this prediction can be regarded as having being satisfied by the data. This difference, however, only occurred in presence of VE relevant priming. This may indicate either that VE experiences are optimal under certain conditions of user priming, or that the presence measures used produce more distinct results under certain conditions of user priming.

10.1.2 Prediction 2 – A difference between the *High stimulus quality/VE relevant priming* and the *High stimulus quality/VE irrelevant priming* conditions

The evidence for this prediction is in Chapter 8, section 8.7.5, and summarized on Table 8.7. The post-hoc analyses of the *High stimulus quality* condition reveals that as measured by the SUS, subjects in the *High stimulus quality/VE relevant priming* group experienced significantly more presence than those in the *High stimulus quality/VE irrelevant priming* group. This finding is repeated in the PQ, but not on the experimental COCI. This prediction can also be regarded as having being satisfied, as it holds for two of the three measures used.

10.1.3 Prediction 3 – No difference between the *Low stimulus quality/VE relevant priming* and the *Low stimulus quality/VE irrelevant priming* conditions

The evidence for this statement is far more obscure. The details are presented in Chapter 8, section 8.7.5 (summary on Table 8.7). The COCI results do not follow the prediction, as the *Low stimulus quality/VE relevant priming* and the *Low stimulus quality/VE irrelevant priming* conditions show a significant difference, with the VE relevant priming imparting an advantage to the subjects in that condition. In the case of the SUS, this pattern is reversed; the significant difference exists, but the VE relevant priming detracts from the presence experienced in the low quality condition. Finally, the PQ shows the predicted pattern – no significant difference between these conditions. With only one of the three measures showing the predicted pattern, it cannot be said, while claiming deference to scientific parsimony, that the prediction has been satisfied by the data.

An item of interest arising from this data, is the SUS result. While there is a difference (which violates the prediction) it is important to note that adding VE relevant priming *decreased* the level of presence experienced by subjects in that condition. This implies that the contribution of conceptual layer activation to presence is not a simple summation or subtraction from the experience of presence (as is the case with the contribution of the perceptual analyzers, which on their own could be modeled by a simple sum). The conceptual layers act in a far more complex way, their effect apparently being to mediate the perceptual context in which the virtual environment is processed.

10.1.4 Considering the evidence graphically

The comparison of the predicted patterns of presence to the empirical findings is better understood by considering the data graphically. Figure 8-10 in chapter 8 presented the predicted presence experiences of the four novel conditions graphically, while Figures 8-4 and 8-5 of chapter 8 represent the means profiles for the measurements made using the SUS and PQ respectively. The COCI results are not used in this analysis, as the COCI failed to match the results of either of the two established scales used. By overlaying these graphs, it is possible to better understand the relationship between these three sets of data more easily.

10.1.5 Displaying the PQ, SUS and predicted scores on one graph

Before this comparison can take place, it is important to note that the sets of data in question use three different scales. The PQ scores have a maximum score of 224; the SUS a maximum of 42, and the predicted scores were not predicted numerically, but rather in a relative way (see section 7.3 in chapter 7 for details). To plot these three data sets onto one graph, it is necessary to transform these scores onto a common scale. We propose the simple method of converting each mean into a proportion of the total possible score on the scale in question for the PQ and SUS scores (these transformations are shown in Table 10-1). Inserting the predicted scores into the graph is more difficult, as only relative scores exist. As no quantities exist for these values, it would be incorrect to compare these directly with the empirically obtained values. Rather, as the purpose of the comparison is to compare the *pattern* of data distribution rather than the magnitude of the scores, the predicted values are simply superimposed over the center of the empirically obtained values to allow the comparison of the data patterns more easily. The resulting graph is presented in Figure 10-1.

Scale	**Stimulus quality**	**Priming**	**Score**	**Maximum possible score**	**Proportion score**
PQ	High	VE relevant	164	224	0.73
PQ	High	VE irrelevant	151	224	0.67
PQ	Low	VE relevant	141	224	0.63
PQ	Low	VE irrelevant	149	224	0.67
SUS	High	VE relevant	29	42	0.69
SUS	High	VE irrelevant	25	42	0.59
SUS	Low	VE relevant	20	42	0.47
SUS	Low	VE irrelevant	25	42	0.60

Table 10-1: Transformation of PQ and SUS mean scores to proportion scores for each of the four conditions

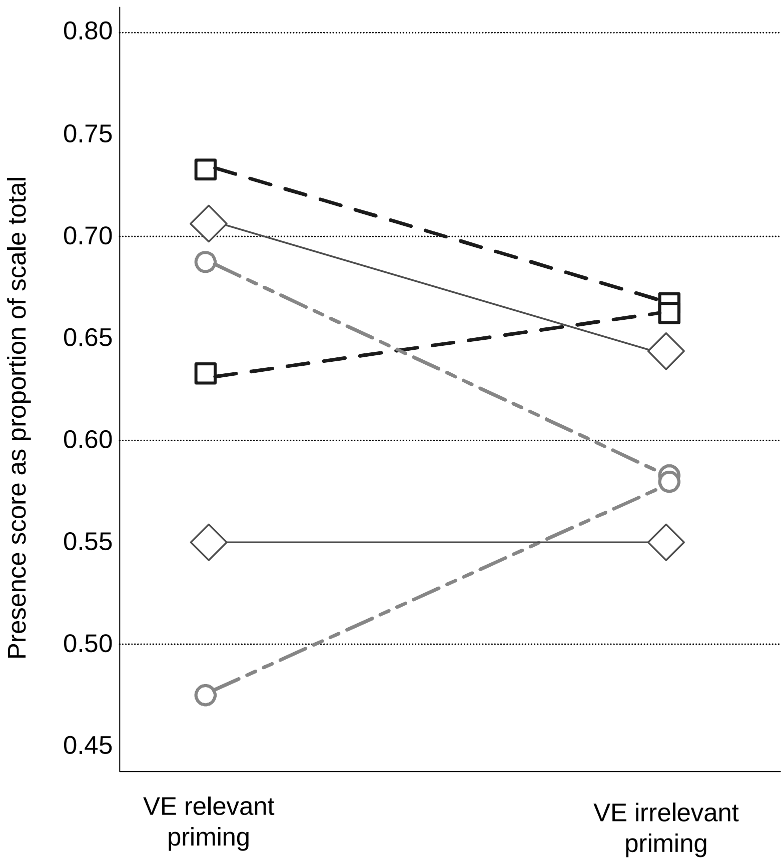


Figure 10-1: The predicted results for the four novel conditions compared to the empirical findings. The blue lines are measurements taken by the PQ, the orange lines measurements made by the SUS, and the green lines the results predicted by the model. For each data set, the upper line shows the *high stimulus quality condition*, and the lower line shows the *low stimulus quality* condition

10.1.6 Examining the data patterns in the graph

With the three datasets superimposed, it is possible to notice that each dataset displays the same basic pattern. Examining the *High stimulus quality* condition for each of the three datasets (the upper line in each case), it is clear that both the PQ (square data points in Figure 10-1) and SUS (circular data points in Figure 10-1) show the pattern predicted by the model – The subjects experiencing VE relevant priming reported more presence than those who experienced VE irrelevant priming. When examining the *low stimulus quality* conditions (lower line in each data set) however, it is obvious that the model's predictions were not realized. The model predicts the same level of presence regardless of priming conditions, and this is indeed the case for the PQ (the slight slope visible is not statistically significant, and can thus be disregarded). The SUS, on the other hand, shows a distinct and significant positive slope, which implies that relevant priming proved detrimental to the presence experience for those in the *low stimulus quality* condition only. As the measurements do not agree in this regard, the evidence must be regarded as contradictory, and for the *low stimulus quality* condition no positive conclusions can be drawn. However, it is possible to conclude a weaker form of the prediction, namely that for the *low stimulus quality* condition, VE relevant priming will *not increase* the presence experienced by subjects.

Another pattern is also discernable in the graph, although it does not form part of the set of predictions made in chapter 7. In the case of both measures, the difference between the *high stimulus quality* and *low stimulus quality* (distance between the two lines in each data set), is marked and statistically significant in the case of *VE relevant* level of priming but insignificant (both in the statistical and literal senses) at the *VE irrelevant* level of priming. The difference between these two priming conditions underscores the role of priming as a mediator rather than as a direct causal agent. This pattern also underscores the importance of priming as a possible source of error variance in experimentation. Figure 10-1 suggests that by controlling the priming state of the subjects, it is possible to exclude mental context as a third variable, and effectively accentuate the difference in presence scores between high and low stimulus quality display conditions.

10.2 Empirical evidence for the independence of O node activation and presence

Section 7.4 in Chapter 7 argued that presence is independent of the perception of objects, and can thus also be achieved through the perception of renditions. To test this notion, a set of expected results was defined in 7.4.1 (shown graphically in Figure 7-11). Experiment 2, described in chapter 9, was performed to test this hypothesis empirically.

10.2.1 Displaying the PQ, SUS and predicted scores on one graph

To better compare the results from the PQ, SUS and the prediction, the three datasets have been displayed together on Figure 10-2. The transformations used to generate a compatible set of scores is described in 10.1.5 above. Table 10-2 presents the raw and transformed scores used to generate the graph (which is displayed in Figure 10-2).

Scale	Display type	Score	Maximum possible score	Proportion score
PQ	Low	140	224	0.63
PQ	Text	120	224	0.54
PQ	High	155	224	0.69
SUS	Low	21	42	0.50
SUS	Text	23	42	0.55
SUS	High	26.8	42	0.64

Table 10-2: Transformation of PQ and SUS mean scores to proportion scores for the three display types

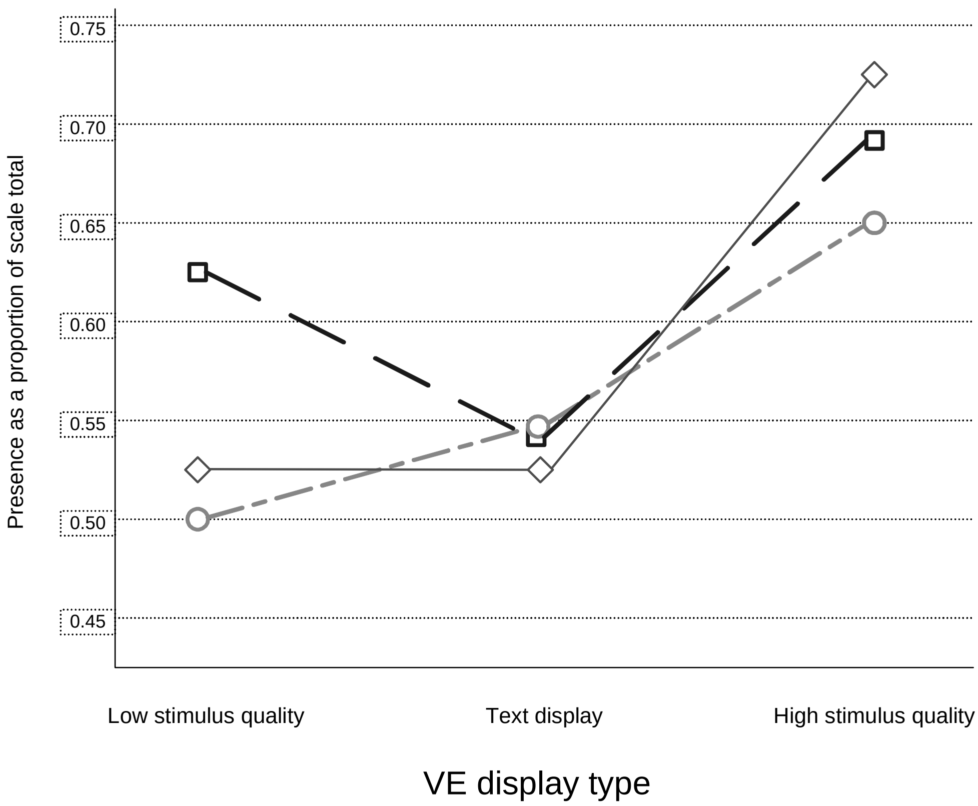


Figure 10-2: The predicted results for the three display types compared to the empirical findings. The blue lines are measurements taken by the PQ, the orange lines measurements made by the SUS, and the green lines the results predicted by the model.

10.2.2 Examining the data patterns in the graph

From Figure 10-2, it is fairly clear that the PQ scores (represented by square data points on the graph) do not fit the predicted pattern (represented by diamonds on the graph). The text display condition scores are lower than the *low stimulus quality* condition, although its score (0.54 of the total PQ score) while being the lowest, is high enough that it could be argued that the subjects were indeed experiencing some form of presence, albeit slight. The pattern displayed by the SUS (represented by circular data points in Figure 10-1) resembles the predicted pattern, although Figure 10-2 shows a difference between the *text display* and *low stimulus quality display* conditions. However, that difference is not statistically significant, (see Table 9-10 in chapter 9) and thus would not be expected to appear in the general case. This means that the SUS follows the pattern predicted by the model.

Sadly, the conclusions which can be drawn about the independence of O node activation and presence are murky. While the SUS shows the predicted pattern, the PQ does not, showing that the perception of objects always lead to increased levels of presence when compared to the perception of renditions. This implies that it is not the virtual environment itself that induces presence, but rather the means used to display it. This effect may be due to the nature of the PQ's items, which weigh the importance of display quality heavily and thus effectively preclude its use in non-graphical displays (this follows the argument presented by Slater, 1999). The PQ might thus not be an appropriate presence measure for use in text-based displays (for a more detailed account of our argument, see section 9.8.3 in Chapter 9). Due to the disagreement between scales, the question of whether rendition biased displays are capable of producing presence remains largely unanswered, although some descriptive results by Towell & Towell (1997) support our prediction that such displays can lead to experiences of presence. Clearly more research is needed to resolve this controversy.

10.3 A critical evaluation of the connectionist model of presence

It is clear from the evidence above that the model presented in Chapter 6 is by no means perfect. The empirical investigations revealed two central deficits. Firstly, the model was unable to predict the negative effect of relevant priming on a low quality display. This is a severe weakness, as this effect was found for users of both the monastery and hospital VEs, which implies that the effect is not an example of an isolated case, but rather a more general phenomenon. If it is the case that this is indeed a phenomenon rather than an experimental artifact, then it is important for the model to be able to recreate this effect. Secondly, the model clearly implies that it is possible for a user report the same degree of presence independently of whether the VE display leads to perceptions of objects or of renditions. The empirical data collected does not reflect this with clarity; a confused mass of results which seem to depend more on the scale with which they were measured than on the conditions under which they were collected, resulted from the experiment created to test the hypothesis.

Although these seem like minor difficulties, these weaknesses are serious because they have the capacity for discrediting the basic structures of the connectionist network. The first problem (i.e. not being able to represent the negative effect of relevant priming on low quality displays) is most serious because it raises the question: *within the constraints defined in the definition of the model, does there exist a set of connections which could represent this effect while maintaining the other properties of the model?* The model used in the empirical tests was not able to show this effect, although it did show a difference in the effect of priming based on the quality level of the stimulus, which is encouraging, as it shows that the model is capable of implementing independence between the effects of priming and display quality. However, it remains to be convincingly demonstrated that the model is capable (or incapable) of replicating this phenomenon. The literature provides very little insights into this problem, as not much published work exists on how a user's previous experience may affect their experience of presence. An brief but interesting idea is presented by Jacobson (2002). Jacobson analyzes the works of literary theorists, and deduces that in extremely low-bandwidth presentations of virtual environments, presence is largely attained by a process of selective self-construction on the part of the user (Jacobson, 2002). The user's mental models of the VE they are immersed in contribute more to the experience than does the actual information received. Jacobson further argues that increasing the amount of information presented to the user will in fact detract from the presence experience, especially when that information does not match with the mental model which the user holds (Jacobson, 2002).

The evidence which we have collected in experiments 1 and 2 (see Chapters 8 and 9 respectively) seem to suggest that Jacobson's hypothesis that more information about the VE leads to lower presence cannot be regarded as a blanket statement which applies to all VE presentation scenarios. Indeed, a large body of work appearing after 1991 (such as that by Sheridan, 1992; Slater, Usoh & Steed, 1994; Witmer & Singer, 1998; Slater & Steed, 2000; and Schubert, Friedmann & Regenbrecht, 2001, to name but a few) which presents both theory and empirical findings to suggest that increasing the amount of information about the VE will improve the presence experience. However, our finding that priming will lead to lower presence levels *in the case of low quality VE displays* does provide empirical evidence for Jacobson's hypothesis.

At first glance, this seems like a contradiction; it cannot be the case that more information about the VE will lead to both higher and lower presence levels. One might also be tempted to disregard Jacobson's hypothesis (together with the evidence we present to support it), on the basis that there exists more published evidence for the idea that providing more information about a VE will lead to higher presence levels. We believe that taking such a step would be an error, for two major reasons. Firstly, the higher proportion of evidence in favour of what might be called the "more information leads to more presence" hypothesis is misleading. Most of the studies which collect such evidence set out from the start to show that more VE information loads lead to higher presence levels; very few have appeared which attempt to show the converse. Thus the overwhelming weight of evidence in favour of the "more information leads to more presence" hypothesis exists, we believe, due to a bias in research direction. Secondly, dismissing Jacobson's hypothesis based on the evidence presented in favour of the "more information leads to more presence" hypothesis is an error because the difference which is being investigated is not only of *information load*, but also of *information display mode*. In the case of Jacobson's hypothesis, the information is presented as text (either spoken of written). In the case of the "more information leads to more presence" hypothesis, the information is presented mostly as moving images. As has been shown by the muddled results arising from experiment 2 (see Chapter 9 for details) and by our arguments against the use of the Presence Questionnaire in text-based VE displays (see section 9.8.3), comparing presence results collected from users experiencing these two types of display cannot be done without running into a host of difficulties.

We propose that Jacobson's hypothesis should not be regarded as a rival to the "more information leads to more presence" hypothesis. Because Jacobson is referring mostly to extremely low bandwidth presentations of VE, we argue, based on our empirical findings, that Jacobson's hypothesis should be applied to the use of low bandwidth displays, while the "more information leads to more presence" hypothesis can be applied to the use of higher bandwidth displays. A cynic might argue that this does not resolve the problem at all, but rather re-states it in a more satisfying form. However, we take the more pragmatic approach and suggest that the use of two hypotheses can be used as a base from which to conduct further research and establish the relationship between the two hypotheses fully.

The second major difficulty with the connectionist model of presence (i.e. the implication that presence can exist in the absence of high quality immersive displays) comes from the fact that R nodes are connected to the action layer, and as such are capable of leading to action node activation. This feature comes from the model's foundation on script theory, as well as the notion that an environmentally suitable behaviour can be triggered by any combination of stimulus and priming state. For instance, an impressionist's painting of a scene would be considered to activate more R nodes than O nodes, but viewing it would still lead to particular thoughts and actions, even if those are different to the thoughts and actions that might occur from viewing a photograph of the same scene. As argued in chapter 3, this is a feature of human behaviour regulation mechanisms, and as such should form a part of any cognition based explanation of presence. The true difficulty in implying that presence can be produced by low quality, non-immersive displays lies not in the cognitive mechanisms involved or in the way that they are modeled, but rather in the way that presence itself is defined. Some authors, such as Winn (1993), include the requirement for an immersive display in their definition of presence, while others, such as Slater (2000), do not explicitly require immersive displays in their definitions of presence.

This difference in presence conceptualization is also reflected in the creation of presence scales. Slater, Usoh & Steed's scale does not make any queries about the system used to display the VE, and it is thus possible for a subject viewing any type of display to score a maximum of 42 points on that scale. The Witmer & Singer Presence Questionnaire, on the other hand, includes items which ask about specific features of the display system. For instance, the PQ item "How well could you identify sounds?" can only achieve a high score if the VE included some form of audio display device; this question has the effect of implicitly lowering the presence score of all subjects viewing a VE without an audio display. It seems that Witmer & Singer have a model or theory of presence causality (which includes the notion that multimodality will always lead to higher levels of presence), and their theory has been expressed in the PQ. This means that the measure will give the most satisfactory results in situations which match the expectations of presence causality held by Witmer & Singer.

In practice this means that subjects viewing high quality multimodal displays will always report high PQ scores, while those viewing text displays will always report lower PQ scores. This difference however, is an artifact of the scale and will occur regardless of the presence experienced by the subjects. This example is a single instance of a larger class of problems in presence measurement, namely the problem of construct validity: no presence scale measures *presence* as a pure concept, but rather *to what extent the subjects' behave in a way similar to what the scale constructor thinks a present subject would behave*. This difficulty exists to some degree in all psychological measurements (Anastasi & Urbina, 1996), but in the case of presence, which has no universally agreed upon conceptualization, the problem is more serious. This problem is not limited to measures of presence. As all work on presence must begin with some definition of the concept, be it implicit or stated, all work which is derived from a particular definition of presence will carry with it the assumptions held by that definition.

The connectionist model presented in this document uses cognitive presence (explained in detail in chapter 4) as its underlying conceptualization of presence. The choice to use cognitive presence was made partly due to the *raison d'?tre* of cognitive presence: it aims to build a bridge between previous major concepts of presence. By combining previous approaches to presence, we hoped to ameliorate the problems described above somewhat. Sadly, fence-sitting in this way comes at a cost. A major drawback to taking a neutral stance is that one is forced to accept the *intersection* between the various presence approaches. This means that a finding can only be regarded as a phenomenon if it exists in all major conceptions of presence; in practice, one is forced to dismiss any result which does not repeat itself in the major presence measures. This conservatism doubtless reduces the available amount of information for building the model because of the risk of dismissing real phenomena as the artifacts of the scales. However, it carries the weighty advantage that all of the phenomena it does include are replicated in several measures, thus making the arguments for their existence as real phenomena more convincing.

10.3.1 A final evaluation of the utility of the connectionist model of presence

Based on the discussion in 10.3 above, it is apparent that creating a model of presence is a difficult task, not only due to the inherent difficulties in modeling complex cognitive phenomena, but also due to the limitations of the contemporary armamentarium of the presence researcher. The model presented here is capable of reproducing many of the phenomena found in the literature, as well as some of those in the experiments designed to test the model. At the same time however, the model shows deficiencies. We suggest that the model should not be considered as an imperfect model displaying several weaknesses; rather, we submit that this is simply the first iteration of a cyclical process which is inevitable in the modeling process. As flaws are uncovered, it is necessary to consider whether the connectionist model is capable of sustaining the modifications required to remove the flaw without straining it to the point of inconsistency. Once the required modifications are made, another series of tests are required to test the changes, as well as ensure that no new faults have been introduced. At the time of writing however, this model presents an empirically tested and theoretically founded method of generating predictions of the relative presence levels which VE users can be expected to experience under particular stimulus and priming conditions.

Chapter 11

Methodological lessons learned

Chapter 10 presented an evaluation of the connectionist model of presence, and a critical discussion of its usefulness in presence research. This chapter looks at some lessons learned regarding the methodology of presence research. We include some conclusions about presence measurement, and some general suggestions for the presence research methodologist.

11.1 Psychometric properties of the Witmer & Singer Presence Questionnaire (PQ)

Psychometric statistics were computed for the PQ as part of the data analysis of experiment 1 (described in Chapter 8). These included several reliability measures and a validity assessment by means of the convergent validity method (Gregory, 1991). The sample size used for this experiment (n=101) was modest for the purposes of determining the psychometric properties of a scale (Anastasi, 1982), and any results found are likely evanescent by comparison to those encountered with a suitably large sample. These analyses are performed simply for the purposes giving the reader some insight into the quality of the scales used.

11.1.1 Internal consistency of the PQ

Section 8.7.6 in Chapter 8 contains the analysis relevant to this section. The internal consistency of a scale refers to the degree to which the items measure a single factor only. Internal consistency is measured by calculating Cronbach's alpha coefficient, which ranges between zero and one; one indicates perfect agreement between items, and zero no agreement between items. The Cronbach's alpha score for the PQ is quite high, being just above 0.9. This indicates that, to a large degree, all the PQ items measure same construct.

A high alpha score is usually an indicator that the scale has only one factor (Cronbach, 1960). However, the PQ contains a number of factors according to its authors. Admittedly, the factors Witmer & Singer refer to are not factors in the statistical sense, but rather in a conceptual sense. These factors include *control*, *sensory*, *distraction* and *realism* (Witmer & Singer 1991). Sadly, the high Cronbach's alpha value implies that all items share covariance, and that there is thus only one factor present in the scale. This data suggests either that all of the factors of the PQ suggested by Witmer & Singer are interrelated, or that their proposition of the existence of these factors is incorrect. This study did not collect enough data to evaluate this question effectively; to do so would require a factor analysis with sample sizes at least double those used in our study (Howell, 2002)

11.1.2 Low reliability items in the PQ

We assessed the reliability of each item in the scale by means of the item-total correlation. This technique determines to what extent a particular item measures the same construct as the rest of the test does. It is calculated by correlating the scores of the item being investigated with the total questionnaire score excluding the item being investigated. Nunnally (1967) suggests that items displaying an item-total correlation of less than 0.4 should be excluded from the scale as they are measuring something other than that which is being measured by the rest of the scale. The item-total correlations obtained for the PQ are shown in table 8-9 in chapter 8.

From that table, it can be seen that a large number of items have adequate item-total correlations, according to Nunnally's criterion. The eight items with low correlations are presented in Table 11-1.

Item number	Item stem
17	How well could you manipulate objects in the virtual environment?
19	How much delay did you experience between your actions and expected outcomes?
20	How quickly did you adjust to the virtual environment experience?
22	How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?
26	To what extent did the events occurring outside the virtual environment distract you from your experience in the virtual environment?
27	Overall, how much did you focus on using the display and control devices instead of the virtual experience and experimental tasks?
29	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?
31	How easily did you adjust to the control devices used to interact with the virtual environment?

Table 11-1: PQ items displaying item-total correlations less than 0.4

An investigation of the items suggests why they produce low item-total correlations. Items 17 and 29 are both directed at the haptic interface to the virtual environment. The version of the DAVE system used in this experiment did not support haptic feedback or other forms of direct manipulation. Consequently, the question *"How well could you manipulate objects in the virtual environment?"* did not apply to the environment at all. This may have caused confusion in the subjects, which in turn would lead to random answering and thus a reduced correlation.

Items 20 and 31 may have a reduced correlation due to ambiguities in the stems themselves. The phrases "how much" and "how easily" do not provide an absolute baseline for the subjects to compare against, and are thus interpreted differently by each subject, which again leads to the introduction of randomness in answering and a reduced correlation. For items 22, 26 and 27, it seems reasonable to expect a large amount of variance naturally (and hence a reduced correlation), as these items rely on the ability to focus attention, which is known to vary greatly form individual to individual (Eysenk & Keane, 1991).

11.1.3 Construct validity of the PQ

The construct validity approach establishes that an instrument is measuring the correct construct by comparing the instrument's performance to other, established scales of the same construct, and by considering the scale's performance in relation to the theory of the construct (Anastasi, 1982). We consider the construct validity of the PQ by using these two methods: by correlating PQ scores with SUS scores, as well as by considering the effect of display quality on PQ scores.

The PQ scores correlate well with SUS scores, achieving a significant $r = 0.76$ in experiment 1 and $r = 0.56$ in experiment 2. An average r for both groups (weighted by sample size) computes to 0.63. Based on these scores, and following the assumption that the SUS measures presence, it can be concluded that the PQ contains a large degree of construct validity, as it agrees to a substantial degree with SUS scores.

According to a large number of presence researchers (for instance Hendrix & Barfield, 1995; Lessiter *et al*, 2001 and Sheridan 1992), the quality of a display will affect the presence experience directly. Specifically, higher degrees of display quality will lead to higher levels of presence. In experiments 1 and 2, presence levels under the *high stimulus quality* viewing condition were compared to presence levels under the *low stimulus quality* viewing condition. The manipulation of display quality led to a significant difference in PQ scores (see Table 8-6 in Chapter 8 and Table 9-9 in chapter 9 for the relevant statistics). This suggests that the PQ is capable of supporting the theoretical underpinnings of the construct.

From the strictly data driven, psychometric perspective, the PQ shows a satisfactory degree of construct validity. However, the evidence presented above relies on a VR display system which matched the conception of display technology held by the authors of the scale. If the PQ is used outside of this domain (as is the case in the text display condition used in experiment 2), the PQ's validity could be called into question (see section 9.8.3 in Chapter 9 for a further discussion of this issue). As discussed in section 10.3 in Chapter 10, this restriction of the domains in which the PQ is applicable should not be considered as a flaw of the scale, but rather as a conscious design decision of which the scale's users need to remain aware.

11.2 Psychometric properties of the Slater, Usoh & Steed presence scale

As part of experiment 1, psychometric statistics were computed for the SUS (these are noted in chapter 8). The statistics and methods are the same used for the PQ (described in section 11.1 above).

11.2.1 Internal consistency of the SUS

The statistics related to this analysis are listed in section 8.7.6 in chapter 8. We expect a high Cronbach's alpha score for this scale, as its authors do not suggest that it is multi-factorial. According to Nunnally, acceptable Cronbach's alpha scores range between 0.8 and 1 (Nunnally, 1967). Values lower than 0.8 are interpreted as revealing a deficit in the reliability of the scale. By this criterion, the SUS shows an unacceptably low Cronbach's alpha score of 0.77. One possible explanation for this lack of reliability might the low number of items in the scale. Anastasi (1982) argues that scale reliability comes from scale effectively sampling the behaviour domain of the construct. Each item in a scale represents a single data point in this sample, and as such, a large number of items are required to sample the behaviour domain adequately and achieve reliability. The SUS contains only six items, which represents a very small sample. Adding items to the SUS might correct this shortcoming.

11.2.2 Low reliability items in the SUS

Using the same procedure described in section 11.1.2, each item was analyzed using an item-total correlation score. The results of this analysis are presented in Table 8-10 (included in chapter 8). Only one item (item 4) displays an unacceptably low correlation according to Nunnally's criterion of 0.4, although item 5 achieves only a borderline score. These two items are presented in Table 11-2.

Item number	Item
4	During the time of the experience, which was strongest on the whole, your sense of being in the virtual environment, or of being elsewhere? *I had a stronger sense of [Being elsewhere][Being in the virtual environment]*
5	Consider your memory of being in the virtual environment. How similar in terms of the *structure of the memory* is this to the structure of the memory of other *places* you have been today? By 'structure of the memory' consider things like the extent to which you have a visual memory of the virtual environment, whether that memory is in colour, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic in your imagination and other such *structural* elements. *I think of the virtual environment as a place in a way similar to other places that I've been to today [Not at all] [Very much so]*

Table 11-2: SUS items displaying item-total correlations less than or approximately 0.4

Examining the items presented in Table 11-2 reveals some possible explanations to why they produce low correlations. Item 4 seems to contain ambiguity brought about by the use of language which is very rich in meaning. For instance the phrase "being in" contains several shades of meaning and as such can be interpreted in various ways by various people. Also, the item forces a dichotomy between "the virtual environment" and "elsewhere". This implies a dual task for the subject. Firstly, the subject has to correctly identify the virtual environment, and then, as a second task, decide the extent of the feeling of being in that place. This can lead to confusion, which in turn leads to random answering and a reduction in the correlation.

With item 5, the reasons for the marginal item-total correlation are likely different. It is noteworthy that this item contains far more words than any of the other items. Item 5 contains 112 words, while the average number of words in the other items is only 31 (with a standard deviation of 3.16). Apart from this, the instruction itself is confusing. It introduces subjects to the concept *structure of memory* and asks subjects to compare one memory to others based on this concept, which is explained only by giving six examples of what constitutes this idea, and suggesting that more properties, which are not explained or defined, exist in this idea. Also, some of the examples given, such as "location in imagination", are anything but clear in their meaning, yet are not explained further. It is quite probable that subjects interpreted this item in various ways (due to the ambiguity in the item), and thus reduced the item-total correlation of this item. If the structure of the VE memory is important to the measurement, as is implied by Slater *et al*, it would have been more salubrious to reliability to reduce the complexity of item 5 by decomposing it into several items, each of which was more clearly explained.

11.2.3 Construct validity of the SUS

To determine the construct validity of the SUS, we correlate the results with those obtained on the PQ, as well as comparing SUS scores obtained by subjects in the *high stimulus quality* condition with the scores obtained by subjects in the *low stimulus quality* condition. This procedure mirrors the one described for the PQ in section 11.1.3 above.

It is necessary to highlight the strategy of using the SUS and PQ to validate each other. Ideally, two scales should not be used to validate each other; a third scale should be used, to create a degree of independence. However, the central purpose of this work is not to unequivocally establish the validity of the scales, but rather to propose and validate the connectionist model of presence. As discussed in 11.1 above, these analyses of psychometric performance are done only to create in the reader some impression of the quality of the measures, and should not be interpreted as rigorous validity analyses of the scales.

The SUS scores display a moderate but satisfactory correlation to the PQ scores (a weighted average *r* value of 0.63 for both experiments, as reported in 11.1.3 above). The SUS also displays appropriate sensitivity to the manipulation of display quality. As presented in Table 8.5 (in chapter 8) and Table 9-7 (in chapter 9), the SUS produces higher scores in the *high stimulus quality* condition. This evidence suggests that the SUS demonstrates a suitable level of construct validity.

11.3 Psychometric properties of the Contents of Consciousness Inventory (COCI)

The COCI is an experimental scale, and as such its psychometric properties are entirely unknown. Although not enough data was collected during this project to convincingly demonstrate its psychometric properties, there is enough data to make preliminary suggestions about the utility of this scale. The determination of the validity and reliability of the COCI were done by the same basic method used in 11.1 and 11.2 above.

11.3.1 Internal consistency of the COCI

Cronbach's alpha coefficient is not suitable for determining the internal consistency of the COCI, because COCI scores are dichotomous (the subject either selected the environmentally keyed word, or they did not). A suitable substitute exists in the Kuder-Richardson 20 (KR20) formula. The KR20 is used as a measure of internal consistency in multiple choice items, and is interpreted in the same way as Cronbach's alpha coefficient (Gregory, 1991). The COCI displays a KR20 value of 0.74 which is low, as acceptable values range between 0.8 and 1. Although, the reliability of the scale is not satisfactory, the value is not so low as to suggest that the COCI is entirely unreliable. As was the case with the SUS, the COCI consists of few items (only ten), and it is possible that the addition of more items might alleviate this problem to some degree.

11.3.2 Low reliability items in the COCI

As in section 11.1.2, each item was analyzed using an item-total correlation score. The results of this analysis are presented in Table 8-11 (included in chapter 8). From this table, it can be seen that a large number of items have low item-total correlations (as per Nunnally's criterion). Five items (items 1, 3, 4, 5, and 6) displayed low correlations. Investigating these items will not lead to any insights, as the

COCI items are simply single words selected randomly. The significant feature of this analysis is the large number of items (half of all the items in the scale) which show poor reliability. This finding, together with the low KR20 value, can be understood to indicate poor reliability in the COCI.

11.3.3 Construct validity of the COCI

The construct validity of the COCI was investigated by the same approach used in 11.1.3 and 11.2.3 above; COCI scores were correlated with PQ and SUS scores, and the COCI's performance was examined under conditions of varying stimulus quality.

The correlations are reported in Table 8.9 in chapter 8 and Table 9-3 in chapter 9. To compute a single correlation across both experiments, a weighted average was calculated to compensate for the difference in sample sizes in each experiment. The weighted average correlation between COCI and PQ across both experiments is r = 0.21, while the weighted average correlation between COCI and SUS across both experiments is r = 0.14. Neither of these correlations is significant, signifying that COCI scores are independent of both the PQ and SUS scores. This does not bode well for the validity of the COCI, as we have demonstrated in 11.1.3 and 11.2.3 above that the PQ and SUS display adequate levels of construct validity.

To further determine the construct validity of the COCI, we examined its sensitivity to changes in display quality, as displayed by the PQ and SUS. It should come as no surprise, based on the poor correlation between the COCI and the other scales reported above, that the COCI performs poorly in detecting this difference. Table 8-8 in chapter 8 and Table 9-4 in chapter 9 show that the COCI was unable to detect any difference in presence levels between the subjects in the *high stimulus quality* condition and those in the *low stimulus quality* condition. The COCI disagrees not only with both of the other two established scales of presence, but also with the current understanding of the concept.

11.3.4 The suitability of the COCI as a presence measure

The evidence presented in 11.3.1 above suggests that the reliability of the COCI is too low to be considered more than a research measure at the moment. By increasing the number of items in the scale, it may be possible to increase its reliability. However, it should be noted that it is not always possible for a measure to show reliability. Cronbach (1960) notes that if the construct which the scale measures displays inherent instability, such as rapid change over time, then the scale cannot be expected to display reliability using the conventional measures. It is not yet clear if presence varies in this way. Slater (2000) suggests that presence probably changes from minute to minute, but offers no empirical evidence. The PQ and SUS are capable of showing high degrees of reliability partly because they are cross-sectional measurements – they sample a single moment in time. The COCI, however, is a longitudinal measure, as it samples at various points throughout the presence experience. This dependence on time will lead to an inherent reduction in reliability. It is therefore possible that temporal stability might not apply as a standard to which the COCI should be held.

The question of construct validity is a different issue. All scales are required to display construct validity (Gregory, 1991). Section 11.3.3 above presents the evidence for the construct validity of the COCI. The low correlation with other presence scales suggests that the COCI measures something other than presence. Also, the COCI is not able to support one of the central tenets of presence theory, namely the effect of display quality on presence. These two lines of evidence suggest that the COCI does not display a great deal of construct validity.

Due mostly to its low construct validity, and partly due to the lack of reliability, the COCI should not be regarded as a useful presence measuring instrument. The basic structure of the COCI could be useful in the future however. Its most useful features include its longitudinal approach, and its avoidance of introspection. The COCI also shows another interesting and useful property. Examining tables 8-4 and 9-4 (in chapters 8 and 9 respectively), which show the effect of priming and stimulus quality on COCI, it can be seen that priming has a reliable and significant effect on COCI scores. Priming leads to a significant difference in COCI scores regardless of the VE which was used, or of the quality of the display. This strongly suggests that the COCI can be used as a measure of priming. We have shown priming to be a determinant of presence, and so the COCI might still have an important role to play in presence research as a priming measure.

11.4 The implications of priming for presence research methodology

The evidence collected while testing the model, discussed in chapter 10, strongly suggests that the role of priming is to set a mental context, which in turn has a strong effect on perception. From this finding it can be inferred that presence is not a direct product of the sensory stimuli presented to a user, but that the mental state of the user plays an important part in the perception of virtual environments and the subsequent selection of behaviour to operate in those environments. This conclusion is partly in contrast to the notion presented by Sheridan, Slater, Witmer & Singer and others that presence is a direct product of the combination of sensory stimulation received by a subject’s modalities. It is true that the presence experienced by a subject depends a great deal upon the stimuli presented to the subject. However, the results presented in chapters 8 and 9 suggest strongly that the mental context in which those stimuli are processed can have a significant effect on the outcome of that processing, and consequently it has an effect on presence as well. The mental context with which users enter the experiment situation is thus a variable which needs to be controlled in an experimental setting, lest it reduce the internal validity of the study by becoming a confound (Cook & Campbell, 1979).

11.4.1 Implications for presence research methodology

If mental context has an effect on the presence experiences, as is argued in 11.4 above, then it is necessary to include a control for this in experiments where presence is used as a dependent variable. Subjects will enter an experimental situation with a variety of mental states, and it is thus not known if the subject’s mental state will increase or decrease their presence experience; this acts as a third variable in the design. One possible solution to this problem is to use random assignment. This will reduce the possibility of unintended priming producing a noticeable effect in presence by ensuring that all groups in the design receive equal numbers of positively primed and negative primed subjects. This solution is not entirely satisfactory, as the great degree of variation in mental states will lead to an increase in error variance, making it harder to detect small effects in the dependent variable (Howell, 2002). This problem can be ameliorated by priming all subjects with the same material, and thereby reducing the overall level of error variance. Although no two subjects will achieve the same mental state from exposure to any particular priming material, the mental states achieved will be similar, and this will lead to the overall reduction in error variance.

11.4.2 Measuring priming

The priming materials used in experiments 1 and 2 (described in chapters 8 and 9 respectively) seem to have been sufficient to induce the primed state and create an effect on the dependent variable. However, it should be regarded as necessary to measure the state of priming to ensure that priming has indeed taken place. Although no measures currently exist for measuring priming, the COCI shows great promise in this regard, as has been argued in section 11.3.4 above. Before the COCI can be used as a measure of priming, however, some developments would be advisable:

1. The low reliability of the COCI should be investigated further. It is necessary to determine if priming is a stable construct, or if it varies from moment to moment. This can be done by adding items to the COCI, and checking if Cronbach's alpha coefficient grows. If it does, then this indicated that the priming construct is stable, and that the scale can be improved to increase its reliability. If not, then the construct is unstable, and reliability should not be regarded as a relevant issue in the measurement of priming.

2. Non-verbal variations of the COCI should be considered. The basic COCI method of selecting from a list on being shown an ambiguous stimulus can be easily applied to images or even sounds. The verbal COCI assumes that the subject understands equally well all the words presented in the list, and as such is not suitable for cross-cultural applications, and replications are not possible beyond language boundaries. A non-verbal variant could help solve this problem.

3. A concerted effort should be made to validate the COCI further. This can be done by creating a manipulation which, according to the theory of priming, should affect the mental state of the subjects. A group of volunteers would then be subjected to the manipulation, and the COCI would be administered. Validation would result if the COCI successfully detected the effect of the manipulation.

Chapter 12

Conclusion

This dissertation presented our explanatory and predictive tool for presence. We developed the concept of cognitive presence, which made use of a cognitive psychology approach in order to combine several strands of presence conceptualization. To measure cognitive presence, we developed the Contents of Consciousness Inventory, a non-introspective measure which is does not use questionnaires, and is administered during the VE experience. We created a cognitive theory to explain and predict our concept, by making use of the connectionist architecture proposed by Martindale (1981) and McClelland & Rumelhart (1986). We then conducted two experiments, each using a large number of participants, to establish the validity of our theory.

12.1 Achievements

In section 1.1 of Chapter one, we provided four stated aims of this project. These were:

1. *To attempt to unify the various strands of presence conceptualization.* We achieved this aim by means of a careful analysis of both the presence literature and the cognitive psychology literature. We created the concept of cognitive presence, and argued that it is a more powerful concept than Lombard & Ditton's *illusion of non-mediation* because we have theoretical evidence that our concept is composed of the atomic processes of environmental processing.

2. *To create a conceptual model to explain and predict cognitive presence.* We made use of the connectionist modeling paradigm to achieve this aim. The three-layer model we created is described in detail in Chapter five. We demonstrated its ability to predict establish presence findings by applying it to the studies by Hendrix & Barfield (1995) and Salln?s (1999).

3. *To establish the validity of our theory by means of empirical investigation.* We designed two experiments to test our model, and predicted the findings of these experiments with the model. We then conducted the experiments with a large group of subjects (103 observations in experiment one and 78 in experiment two). A summary of the findings of these experiments can be found in 12.2 below. These experiments largely confirmed the main thrust of our theoretical concepts.

4. *To create a measure of cognitive presence.* We achieved this aim partially. Although we successfully created the Contents of Consciousness Inventory, its use in the field indicated that it is not a good measure of presence (see section 11.3 of Chapter 11 for a discussion of this point). However , we were still able to verify our concepts by means of established measures of presence.

12.2 Summary of empirical results

The two experiments we conducted provided a variety of data to evaluate our main hypotheses as well as provide some interesting ancillary information such as the psychometric properties of the presence measures we used.

12.2.1 The effect of display quality on presence

Between both experiments, we tested the effect of three levels of display quality on presence, as measured on three presence scales: the Slater, Usoh & Steed scale (SUS), Witmer & Singer's Presence Questionnaire (PQ), and our own Contents of Consciousness Inventory (COCI). The three quality conditions were:

1. High-graphical quality display, which used radiosity, texture mapping, and positional sound

2. Low quality graphical display, which used Lambertian lighting, no texture mapping and no sound

3. Text-based display which used text descriptions of rooms and non-realtime interaction.

We found that the COCI did not respond to changes in display quality. We also found that the high quality graphical display consistently produced the highest levels of presence on both the SUS and PQ. The text condition produced SUS scores which lay between the high and low quality graphical displays, although on the PQ, the text display consistently produced the lowest scores. These findings (which include the appropriately primed subjects) are in line with the bulk of the presence literature, which suggest that the level of immersion of a VE system directly affects presence. However, if one considers only the data which were obtained from subjects which were not appropriately primed, then no difference in display quality was apparent. This implies that priming plays an important role in facilitating presence differences under conditions of varying display quality.

12.2.2 The effect of priming on presence

We used two priming conditions in our experiments. These were:

1. VE relevant priming: before entering the VE, the subject read through a booklet which was related to the theme of the virtual environment

2. VE irrelevant priming: before entering the VE, the subject read a booklet unrelated to the theme of the virtual environment

As with the results quoted in 12.2.1, the COCI failed to respond to the manipulations; the PQ and SUS, however, both responded significantly. We found no main effect with regards to priming and presence; that is, we found that priming condition did not directly cause a change in presence scores. However, we did find a significant interaction effect between stimulus quality and presence. In the *VE irrelevant priming* condition, there was no difference in presence levels between display quality conditions. In the *VE relevant* priming condition, however, there was a difference. Furthermore, the relevant priming increased the score for those in the high quality display condition, and lowered the score for those in the low quality display condition.

12.2.3 Psychometric evaluation of the presence measures

We conducted reliability and validity analyses on all three presence measures used. We used three criteria for each measure, namely:

1. Cronbach's alpha coefficient as a measure of internal consistency (the Kuder-Richardson 20 formula was used for the COCI, as this uses a hit/miss scoring scheme)

2. Correlations with other presence scales to determine concurrent validity

3. Examination of sensitivity to display quality manipulations to determine construct validity

For the SUS, the Cronbach's alpha score was unacceptably low, probably due to the low number of items in that scale. The SUS did show a significant, positive correlation to the PQ, suggesting that it measures the same construct to a degree. Finally, the SUS was able to distinguish the difference between presence scores in conditions of differing display quality, as suggested by Slater, Usoh & Steed (1995). We conclude that the SUS is capable of measuring presence, although its reliability is in question.

The PQ also performed well. It showed a satisfactorily high Cronbach's alpha score, suggesting that it is internally consistent. However, this is a concern due to the fact that a high Cronbach's alpha score suggests the scale measures only one factor, while Witer & Singer state that the PQ in fact measures several factors. The PQ correlated well with the SUS, suggesting that it is suitable to measure presence. Finally, the PQ was able to distinguish presence scores from two differing display conditions, in line with its authors' expectations (Witmer & Singer, 1998).

The COCI did not perform well. It showed a low Kuder-Richardson 20 score, suggesting that the items are not measuring the same construct. The COCI also failed to correlate with either the PQ or the SUS, suggesting that it is not measuring the same construct as these scales. Finally, the COCI was not able to distinguish between presence scores in differing display quality conditions, as is expected for a presence measure.

12.3 Summary of the validation of the connectionist model

After completing the experiments, we considered if the evidence collected supported our model. We did this by comparing the obtained results, with the results predicted by the model. The first experiment aimed to validate the basic structure of the model. For the first experiment we made three predictions:

1. A difference between the *High stimulus quality* and *Low stimulus quality* conditions: We found a significant difference on two of the three presence measures.

2. A difference between the *High stimulus quality/VE relevant priming* and the *High stimulus quality/VE irrelevant priming* conditions: We found a significant difference on two of the three presence measures.

3. No difference between the *Low stimulus quality/VE relevant priming* and the *Low stimulus quality/VE irrelevant priming* conditions: We found no difference with one scale (the PQ), and a difference on another (the SUS). The difference created a disadvantage for those subjects in the *Low stimulus quality/VE relevant priming* condition.

Two of the three predictions were supported by the data, with the third being partially supported. We conclude that this experiment provides a substantial degree of support for our model. The second experiment aimed to demonstrate that presence is independent of the medium used to display the environment. We generated one prediction with the model; that there would be no difference in presence scores collected from a group viewing a text-based display and a group viewing a low-quality graphical display, while a group viewing a high-quality graphical display would fare better than both the low-quality graphics group and text group. The data collected were somewhat contradictory. The SUS displayed the predicted pattern, but the PQ did not. This hypothesis is thus insufficiently supported.

12.4 Contributions of this work

This dissertation contains several novel contributions. The major theoretical contributions of this dissertation are:

1. An attempt to unify the various strands of presence conceptualization by redefining them in terms of well-understood psychological processes which ensure environmentally appropriate behaviour. We call our concept *cognitive presence.*

2. A strategy to measure cognitive presence. This is does not require questionnaires or introspection, and is administered during the VE experience to minimize memory distortions. We call our measure the *Contents of Consciousness Inventory* (COCI).

3. A conceptual model to explain and predict cognitive presence. This model makes use of a connectionist architecture as defined by McClelland & Rumelhart (1986). The model includes both top-down and bottom-up processes. This is, we believe, the first model of its kind to give equal importance to these two fundamental processes.

4. The discovery of priming as a major determinant of the presence experience. Of particular interest is the finding that priming does not act as a simple effect, but rather interacts with display fidelity.

To determine the validity of our theoretical analyses and conclusions, we conducted two experiments, making use of a large data set (103 observations in the first experiment, and 78 in the second). These experiments resulted in a number of empirical contributions:

1. We found that *priming* (the manipulation of top-down processing) has an interesting and powerful effect on presence. Those subjects who were primed to respond to positively to the VE responded with higher presence score *only if they were viewing the VE with a high quality display*. However, if the quality of the display was low, *positively primed subjects reported less presence than those who were not primed for that VE*. Those subjects who were not primed to respond to the environment reported the same level of presence, regardless of the quality of the display. This finding demonstrates not only the power of priming to change presence levels, but also shows that detecting presence differences among improperly primed subjects could lead to difficulties.

2. We found that, ignoring priming, better quality VE displays lead to more presence. This finding replicates the findings by many previous researchers, such as Hendrix & Barfield (1995), Slater, Usoh & Steed (1994) and Salln?s (1999).

3. We found that text-based VE displays were generally able to produce presence on a par to that produced by low-quality graphical displays.

4. We performed psychometric analyses of the presence scales used. We found that Slater, Usoh & Steed's scale (1995) showed good validity, but poor reliability. Witmer & Singer's Presence Questionnaire (1998) showed both good reliability and good validity. Finally, our COCI scale showed poor validity and poor reliability.

12.5 The discovery of *priming* as a mediator variable in presence

Perhaps the most striking (and serendipitous) finding of this work is the identification of the effects of subject priming on their experiences of presence. Originally, we expected priming to act as an extra source of information for subjects, and so we expected it to behave simply as a causal agent (see section 7.3 in Chapter seven for a detailed discussion of our expectations in this regard). However, the empirical findings show distinctly that priming can act both to increase or decrease presence, depending on the quality of the VE display. This finding is unlikely to be a simple experimental artifact, as it appeared in both the monastery and hospital VEs, and appeared in measurements taken with both Slater, Usoh & Steed's scale, and with Witmer & Singer's Presence Questionnaire.

Priming is particularly interesting for two reasons. Firstly, priming is created not by "preparing" the subjects for the VE experience by telling them what they will be expecting. Its effect can thus not be explained by a simple case of users being disappointed at not finding what they were told about in the preparation session The priming manipulation is far more subtle, and has only to refer to the *theme* of the VE in order to be effective. It can thus be best explained as creating a bias in the perceptual processes of the subjects before they enter the VE. Secondly, priming is interesting because it is a *mediator* to presence rather than a direct cause. If one maintains display quality constant, and manipulates priming, no effect is discernible, because it is not a cause. It operates differentially on different levels of a cause variable (in this case, display quality), and can thus only be detected by means of a factorial design. As our results show, although it is not a direct cause of presence, it is still capable of leading to noticeable and statistically significant differences in presence scores.

The role of top-down processing effects on presence have not been adequately investigated, perhaps due to the supremacy of Gibsonian thought in the presence community (see Zahoric & Jenison, 1998 and Sheridan, 1999 for examples of this type of work). Gibson gave paramount importance to bottom-up processes in perception, due to his belief that all of the information necessary to operate in an environment was available via the senses. This view of perception is not largely adhered to by cognitive psychologists, following the demonstrations of the power of top-down processing by Pollack & Pickett (1964), Reicher (1969), Williams & Weisstein (1978), and many others. Based on our finding, it would seem that Gibson's ideas might also be of less use to presence research than was once believed.

12.6 Future development of the connectionist model of presence

The model described in this document presents many opportunities for improvement and further investigation. This section presents some directions of research we feel should be followed as a first step towards improving the model. We present these suggestions in an order which we believe represents their relative importance, beginning with the most pressing.

12.6.1 Establishing the independence of O node activation and presence

As noted in chapter 10 (in section 10.2), the two experiments performed failed to clearly demonstrate the independence of presence from O node activation. This issue needs to be addressed with some urgency, as the structure of the perceptual analyzers is created under the assumption that O and R node activation can be equivalent under the right conditions.

To establish that presence can arise as a result of either O or R node activation, it would be necessary to replicate the basic design of experiment 1, but using several levels of priming and stimulus quality rather than simply two levels of each. Particularly, it would be necessary to widen the range of each variable, to include more extreme cases. For example, for the *stimulus quality* variable, immersive devices such as head-mounted displays could be used to maximize the degree of O node activation, while priming could be maximized by the use of multimedia shows rather than simple text. Using extreme levels of these variables can be helpful, because this allows one to explore the possibility that the independence occurs only at very high or very low levels of stimulus quality.

Based on the findings of such research, it might become necessary to amend the model. If it is found that R nodes are not capable of contributing to presence, or that they can only contribute under extreme conditions, the necessary changes might be effected by changing the strength of the connections between the R nodes and the action layer. If the R nodes do not contribute at all to presence, this might be modeled by connecting all R nodes to a small subset of action nodes which represent behaviours that are usually unsuitable for operating in an environment. If it is found that R nodes only contribute to presence under extreme conditions (e.g. under high levels of R node activation only), this could be modeled by reducing the strength of the connections between the R nodes and the action layer.

12.6.2 Decomposition of the conceptual layer into constituent analyzers

As described in section 5.2 (in chapter 5), the conceptual layers in the current model represent a more complex structure, which was not elaborated on due to a lack of relevant empirical evidence. Modeling the contribution of mental state to presence by means of a single source of priming does not allow for an understanding of how different elements of cognition, such mental models of environments and previous experience, each contributes to presence. Once some components of the conceptual layer have been identified, it might be possible to better understand the negative effect of priming on presence in low quality displays. Knowing the components which make up the conceptual layers will allow for the establishment of a priming theory, which can be used to determine the appropriate type of priming to maximize presence for a particular perceptual situation.

12.6.3 Creation of an action layer activation measure

The COCI was conceived as a measure of cognitive presence, and the connectionist model mapped cognitive presence onto the activation of specific action nodes. However, as discussed in section 11.3 of Chapter 11, the COCI lacks the required degree of construct validity. This leaves the requirement for a measure of cognitive presence unfilled.

Although the other presence measures used in this study lead to satisfactory results, they should not be regarded as suitable measures of cognitive presence or of action layer activation. The success of these other measures in this role comes from the fact that action layer activation is *related to* but not *synonymous with* traditional conceptions of presence. These measures doubtless include aspects of their own conceptions of presence which are not related to action layer activation (such as the measures of display system features in the PQ). These can act as confounds and reduce reliability when used to measure action layer activation. It is necessary to search again for a measure of action layer activation.

As refinement of the connectionist model is by necessity driven by empirical research, it is important to have an accurate, direct measure of the degree to which experimental manipulations are affecting the model. The basic concept behind the COCI still merits investigation to determine if it is a useful method for measuring action layer activation.

12.6.4 Maturing the COCI into a dedicated priming measure

As argued in section 11.3.3 and 11.3.4 of Chapter eleven, with some refinements the COCI could be used as a measure of priming. The current version of the COCI is already quite effective, but it can doubtless be improved. The major flaw of the current COCI seems to be its dependence on the content of the VE on which it is to be used, effectively precluding the creation of a "universal COCI" to work on any environment. Based on the theory of priming, it seems unlikely that priming can be measured without making reference to the content of the mental state, as priming is defined in terms of content itself; to say "the subject is primed" is really a contraction of "the subject is primed for *x*", and as such measuring priming without reference to *x* is impossible. Requiring a new version of the COCI for each environment can however create a serious confound, as each version of the COCI might posses differing psychometric qualities.

This problem can be partially circumvented by establishing a rigorous method for creating COCI variants. This method would aim to reduce possible error variation associated with the choice of particular items over other possible candidate. One solution to this already exists in a related problem in psychometrics, namely the *parallel forms problem*. The problem of parallel forms is to create two separate tests of the same subject, with equivalent psychometric properties, and of equal difficulty. The standard solution to this problem is to create one test of twice the length, and to use a random selection method to select items for each of the two tests. A similar procedure could be applied to the COCI – rather than selecting ten words for the test, a longer list is created, and ten words are chosen at random from the list. As it is likely that only a subset of all words will be problematic, this method reduces the probability that problematic words will be selected for inclusion in the test.

12.6.5 Find connection strengths and unit functions

Before a connectionist network can be used to make quantitative predictions, the strengths of the connections and the unit functions must be determined (Martindale, 1990). By way of assumption, the model currently treats all connections as if they were of equal strength. However, it is unlikely that this is the case, due to the effects of learning and experience on connection strength.

An advantage of working with a model that specifies the connection strengths is that it makes it possible to incorporate individual differences into the model, which allows the modeling of particular networks (i.e. particular people). To successfully predict the presence that a particular individual will experience, an "average network" would be used, which models the mean performance of subjects. Then, some method of quantifying the difference between the individual in question and the average network (such as, for instance, Witmer & Singer's Immersive Tendencies Questionnaire) can be used to calibrate the network for the individual in question.

To determine connection strengths, an accurate measure of each construct attached to a particular connection is required. For example, to determine the connection strength between the O node of the visual analyzer and the action node encoding *visiting a restaurant*, a measure of each is required. After activating one of the nodes to varying levels, the activation of the other connected nodes can provide the required information to determine the strength of the connection as well as the parameters of the unit function.

12.6.6 Build network simulator for automatic prediction

Once accurate measures exist and the network parameters have been calculated, it will be possible to implement a simulation of the network. This will allow a quick method of predicting presence quantitatively for a given set of initial conditions. A simulator can be a useful tool for designers of virtual environments, as it would allow them to determine quantitatively what effect a change in the VE or priming manipulation would have on the presence experienced by users in that environment.

A network simulator is also a useful tool for research purposes, as it allows detailed examinations of the models. With detailed analyses possible, researchers working in applied fields such as VR therapy and collaborative shared spaces could quantitatively investigate the relative benefits of presence. Also, as detailed analyses of how activation spreads to the conceptual layers would become possible, it would become feasible to analyze the possible contribution of presence to performing a particular task, given that one knows the cognitive components required to complete the task.

Appendix A

The DAVE tool

We developed the DAVE tool to support the experiments described in this dissertation. DAVE is an acronym which stands for *Dave's Application for Virtual reality Experiments.* DAVE runs on the Win32 platform, and has been tested and found to work on Windows 98 sp2, Windows 2000 sp 1 and Windows XP. DAVE requires DirectX 7 or later to be installed. Although it is not a requirement, a great increase in performance is obtained from using a Direct3D compatible graphics accelerator. We used version 1.1 of the Genesis 3D SDK, and compiled DAVE using Microsoft Visual C++ version 6.0 sp 1.

Figure A-1: An example of the rendering capabilities of DAVE. This scene can be rendered with a refresh rate of 20Hz on an AMD Athlon 700MHz with a GeForce 2MX based graphics card.

A.1 Graphical rendering capabilities

At DAVE's core is the Genesis 3D game engine (http://www.genesis3d.com). Genesis 3D was written by WildTangent Inc, and released to the public under an open source license in 1998. Genesis 3D has the following central features:

- Texture mapping
- Soft-edged shadows
- Soft scene lighting by means of the radiosity method
- Portal based occlusion culling
- Fog

DAVE uses these features to implement a scene visualization tool, which rendered the scene from a first-person perspective, by displaying a camera which is set 1.7 virtual meters from the floor of the scene, to give the illusion of standing in the scene (see Figure A-1 for example). During the experiments, the screen refresh rate of DAVE did not fall below 9Hz, and was typically at 18Hz or above.

A.2 Sound rendering capabilities

The Genesis 3D API provides for stereo sound rendering, and the dynamic modification of sounds. DAVE uses this feature to implement two types of sound: ambience sounds and position sounds. Ambience sounds are used as aural "filler" for the scene. For example, the medieval monastery scene contained the sounds of birds singing, wind blowing and distant thunder as ambience sounds. Ambience sounds are chosen randomly from a list provided by the scene designer, and are played with random settings for volume, pan and frequency. Position sounds are inserted into particular points in the scene by the scene designer. DAVE uses the pan and volume controls provided by Genesis 3D to simulate distance attenuation and sound position. This gives the illusion that position sounds are being emitted from particular places in the scene.

A.3 User interaction and navigation

DAVE allows the users to move around the world by means of the *quake keys* navigation method (Dalgarno & Scott, 2000). This method uses the mouse and keyboard to navigate. The mouse is used to change the yaw and pitch parameters of the camera (the mouse x axis controlling yaw, and the y axis controlling pitch), while the keyboard is used to move in relation to the camera's view vector (the W and S keys move forwards and backwards along the view vector, while the A and D keys move left and right perpendicular to the view vector). As the camera position changes to these inputs, the height of the camera is kept at a constant height above the floor of the scene at that point, to create the illusion of walking over the scene. The version of DAVE used to display the scenes used for the experiments described in this dissertation used a 17" monitor to display the scenes, although we have subsequently modified DAVE to display stereo images through the Virtual Research (http://www.virtual research.com) VR6 head-mounted display.

A.4 "Hunt the tokens" game

DAVE also implements a simple game. The scene designer has the option of placing a number of waypoints (referred to as "tokens") in the scene. The aim of the game is to traverse the path defined by these waypoints. In the display, the tokens are displayed as textured brown boxes. DAVE only displays one such box at a time; it is placed at the co-ordinates of the next waypoint along the route. Once the camera's position comes within a particular threshold (1 virtual meter) of the waypoint, DAVE plays a sound, and sets the current waypoint to the next along the route. The box is then displayed at the position of the next waypoint. This gives the illusion of “picking up” each box along the route.

To aid the user to find the tokens, DAVE displays a small compass-style device (which we refer to as the “scanner”) on the bottom right corner of the display. This shows the relative bearing to the next waypoint, as well as an indication of the token’s relative height by means of three states: above/level/below (Figure A-2 shows the scanner at work). The scanner is also capable if indicating range to the waypoint, by means of a sound signal. To access this feature, the user clicks the left mouse button. While the button is held down, a continuous sound is played, and the pitch of this sound indicates the relative distance to waypoint. A low frequency indicates a long distance, a high frequency a short distance.

Figure A-2: The scanner used for the "Hunt the Tokens" game can be seen in the bottom right of the DAVE display

A.5 Interaction recording and playback

DAVE allows for the optional recording of a user’s interaction with the VE, and the subsequent playback of this record. This is accomplished by a time sampling of the camera’s position and orientation at regular time intervals. These samples are then buffered, and periodically written to the disk. With a sampling frequency of 10Hz and a buffer size of 500 samples, the recording mechanism does not introduce any noticeable lag into the system. To play the record back, DAVE is started in a special mode, which reads the record and linearly interpolates the samples to create a smooth, real-time playback of the user’s interaction. This allows for careful examinations of the user’s behaviour in the scene.

A.6 COCI scale implementation

DAVE also implements the COCI presence measure described in Chapter four of this dissertation. The scene designer has the option to include a set of COCI items (of the verbal variety) to display to the user. The COCI engine works as a simple finite state machine. At a certain time interval, which has a small random component, DAVE enters the COCI display state. It then works through the following states:

0. Play the COCI warning sound to the user
1. Clear the screen and wait 0.2 seconds
2. Display the word fragment for 0.2 seconds
3. Clear the screen and wait 0.2 seconds
4. Present the user the word list (see Figure A-3 for an example list). If state 4 has lasted more than 25 seconds, mark this COCI item as unanswered, and return to the scene display/interaction state. If the user presses keys F1, F2, F3 or F4, record the key pressed and return to the scene display/interaction state.

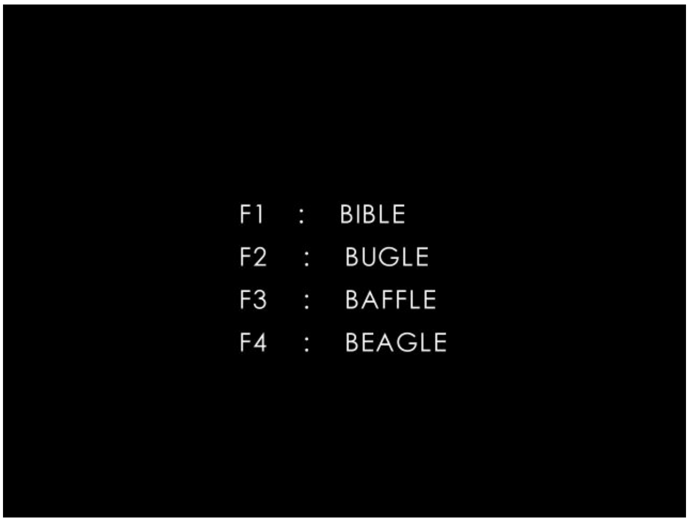


Figure A-3: DAVE displaying the COCI word list

DAVE saves the COCI results file on one of two events. The first is the user pressing the exit key (F12). The second is if the last COCI item has been displayed (i.e. state 4 above has completed). This ensures that even if a user did not complete the entire set of 10 COCI items, the items which they did respond to would be available for analysis later. The responses are tagged with the date and time as well as a single letter identification code which is set in a configuration file. This was done to allow the COCI responses to be matched up to the pen-and-paper questionnaires used.

Appendix B

The TIVE tool

We developed the TIVE tool to support the second experiment described in this dissertation. TIVE is an acronym which stands for *Text Interface to Virtual Environments.* TIVE runs on the Win32 platform, and has been tested and found to work on Windows 98 sp2, Windows 2000 sp 1 and Windows XP. TIVE requires DirectX 5 or later to be installed. We used version 4.0 of the Allegro library, and compiled TIVE using Microsoft Visual C++ version 6.0 sp 1.

B.1 Rendering capabilities

TIVE was implemented using the Allegro 4.0 game programming library which was written and released under the open source license by Shawn Hargreaves (http://www.talula.demon.co.uk/allegro). TIVE renders still images and text onto a 320x240x16 display. TIVE does not support sound rendering.

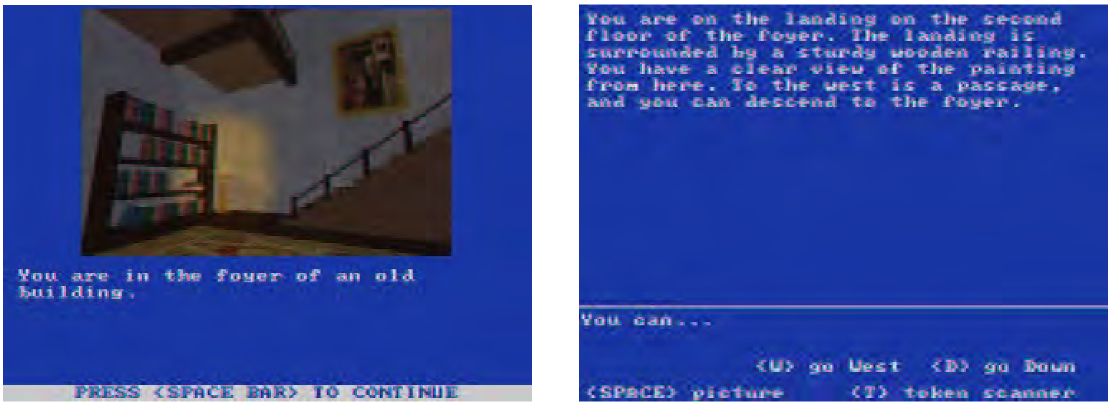


Figure B-1: The TIVE display; Still image on the left, word description on the right (the space bar toggles these two displays)

B.2 User interaction and navigation

TIVE is a non-realtime display engine. Updates of the engines are initiated only as a response to user input events. TIVE displays a single “room” at a time, by means of either a text description only, or a text description accompanied by a still image. Included in the display is the set of exits of the current room into other rooms. There can be a maximum of six exits from any room: North, South, East, West, Up and Down. The user is only shown the exits which the room actually contains (i.e. if a room does not have a North exit, then North will not appear in the list of exits). To use an exit, the user presses a single key – N for North, S for South, E for East, W for West, U for Up or D for Down. To help the user remember the interface, these keys appear on the display. See Figure B-1 for an example.

B.3 "Hunt the tokens" game

TIVE also implements a simple game. The scene designer has the option of placing a number of waypoints (referred to as "tokens") in rooms in the environment. The aim of the game is to traverse the path defined by these waypoints. The tokens are not displayed, but the user is notified on entering a room which contains a token. TIVE only informs the user of one token at a time; it is placed at the co-ordinates of the next room along the route. When the user enters the room with the token, a window appears informing then that the token has been found, and the current waypoint is set to the next along the route. This gives the illusion of "picking up" each token along the route.

To aid the player to find the tokens, TIVE can display a window informing the user of the bearing (using compass directions) and relative height of the next token (this window is referred to as the "scanner"). To access this feature, the player can press the T key at any time. Figure B-2 shows the scanner display.

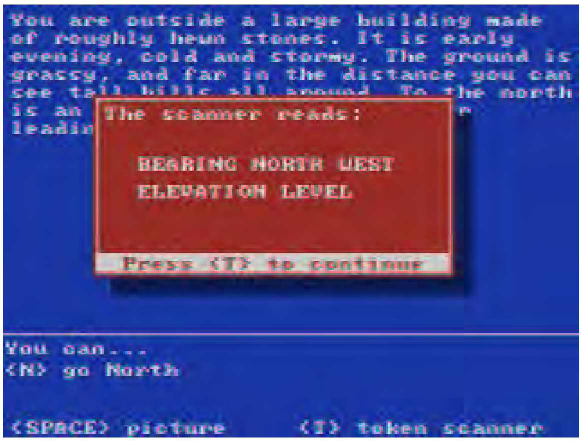


Figure B-2: The scanner in TIVE showing the direction to the next token

B.4 Interaction recording and playback

TIVE allows for the optional recording of a user's interaction, and the subsequent playback of this record. This is accomplished by recording each keypress made by the player. To play the record back, TIVE is started in a special mode, which reads the record and executes the next keypress every fifteen seconds. This allows for careful examinations of the user's behaviour in the scene.

B.5 COCI scale implementation

TIVE implements the COCI presence measure described in Chapter four of this dissertation. The scene designer has the option to include a set of COCI items (of the verbal variety) to display to the user. The COCI engine works as a simple finite state machine. At a certain time interval, which has a small random component, TIVE enters the COCI display state. It then works through the following states:

1. Clear the screen and wait 0.2 seconds
2. Display the word fragment for 0.2 seconds
3. Clear the screen and wait 0.2 seconds
4. Present the user the word list (see Figure B-3 for an example). If state 4 has lasted more than 25 seconds, mark this COCI item as unanswered, and return to the scene display/interaction state. If the user presses keys F1, F2, F3 or F4, record the key pressed and return to the scene display/interaction state.

Figure B-3: TIVE displaying the COCI word list

TIVE saves the COCI results file on either of two events. The first is the user pressing the exit key (Shift-Q). The second is if the last COCI item has been displayed (i.e. step 4 above has been completed). This ensures that even if a user did not complete the entire set of 10 COCI items, the items which they did respond to would be available for analysis later. The responses are tagged with the date and time as well as a single letter identification code which is set in a configuration file. This was done to allow the COCI responses to be matched up to the pen-and-paper questionnaires used.

Appendix C

Images from the Hospital VE

This appendix shows eight images taken from the high and low stimulus quality versions of the Hospital VE used in the experiments described in this dissertation. The VE is not based on any real place. In each case, the left-hand image is the high stimulus quality version, and the right-hand image is from the low stimulus quality version. Apart from the graphical differences, the high stimulus quality version included both positional and ambient sounds, while the low stimulus quality version did not include any sound.

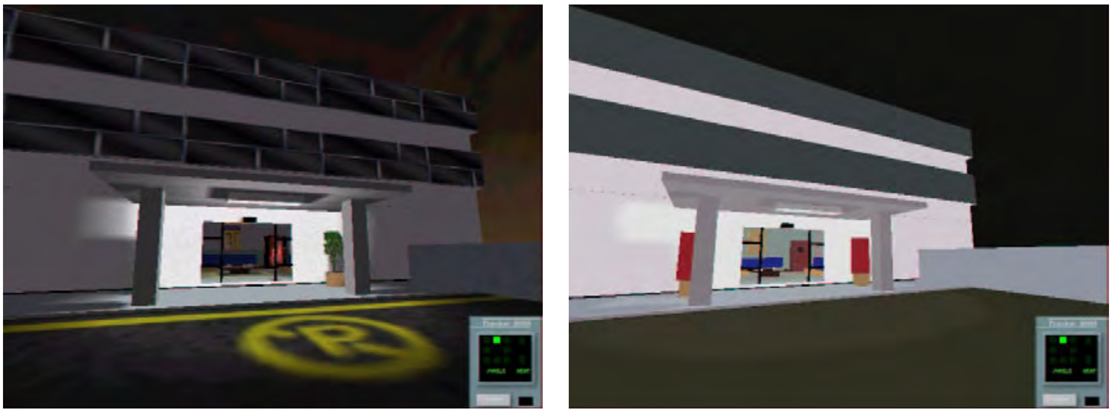


Figure C-1: The outside of the hospital

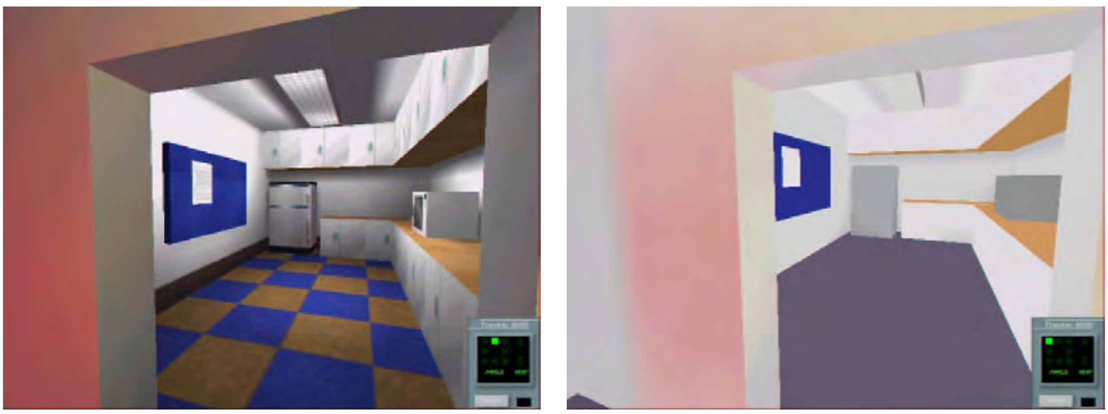


Figure C-2: The kitchen of the staff room

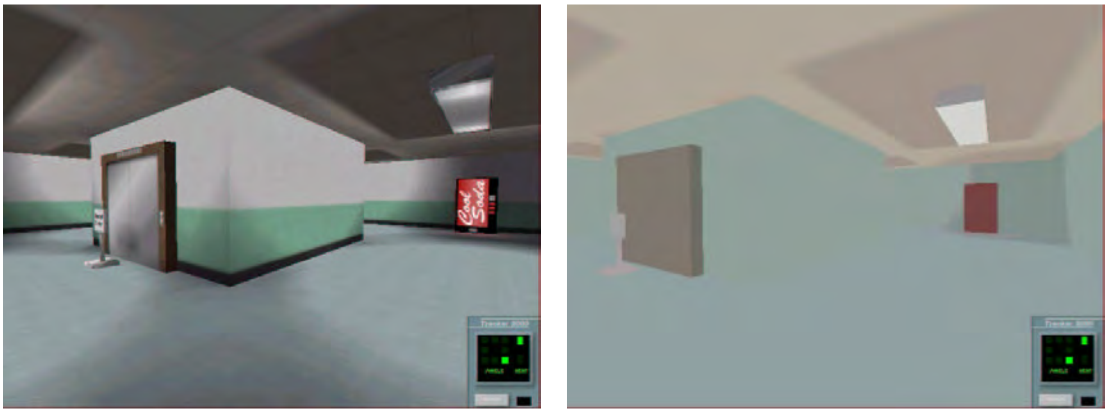


Figure C-3: Elevators and vending machine (basement, level 0)

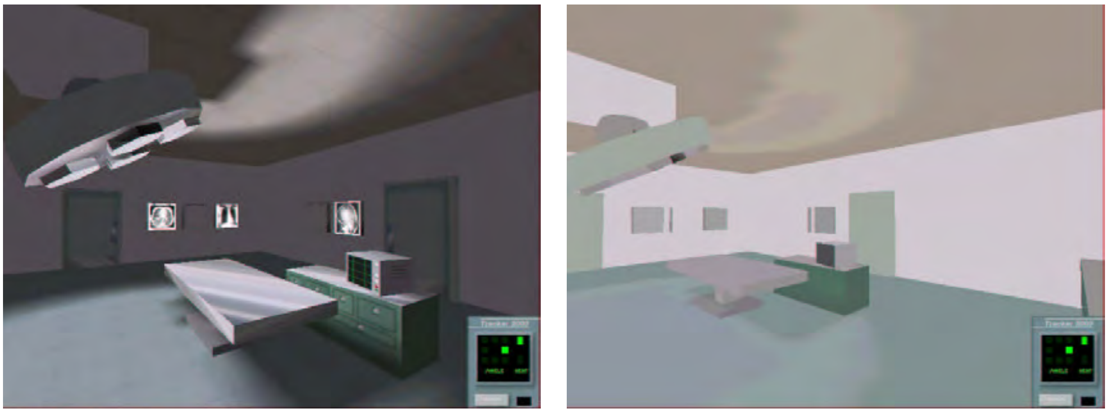


Figure C-4: Surgical Theater

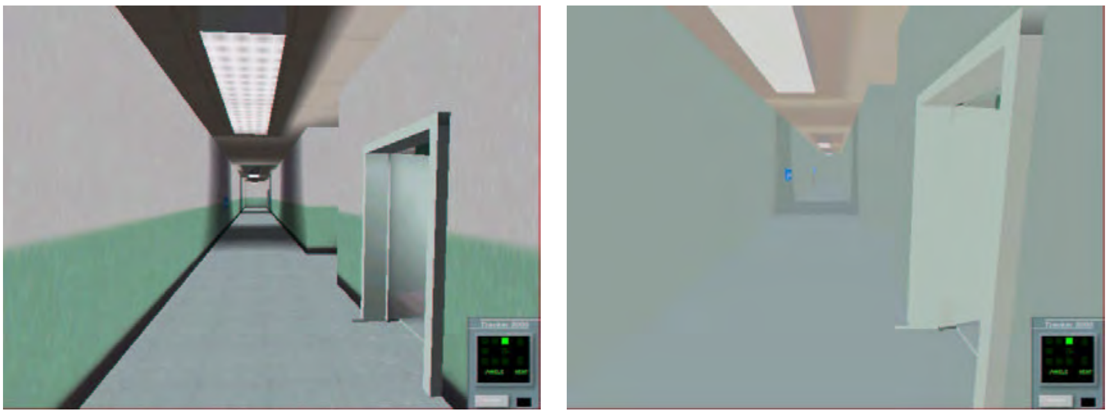


Figure C-5: Passages (ground floor, level 1)

Figure C-6: Eastern stairwell (first floor, level 2)

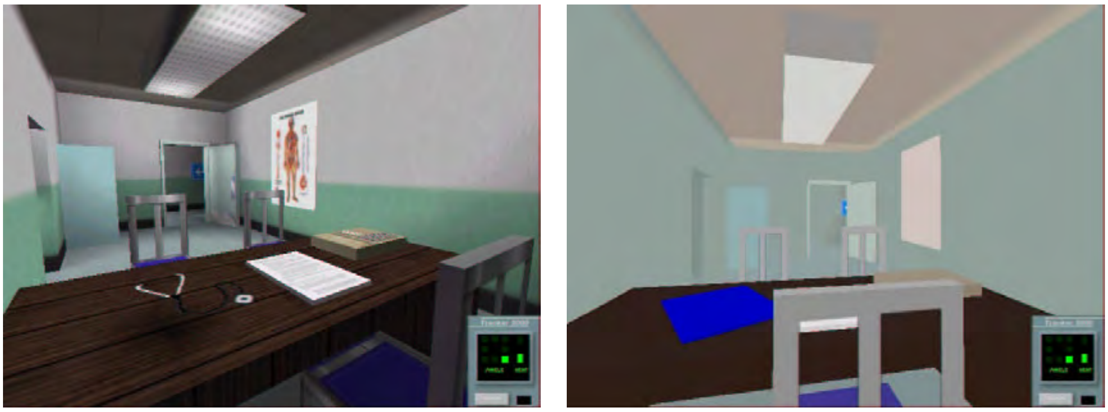


Figure C-7: Consulting room (level 3)

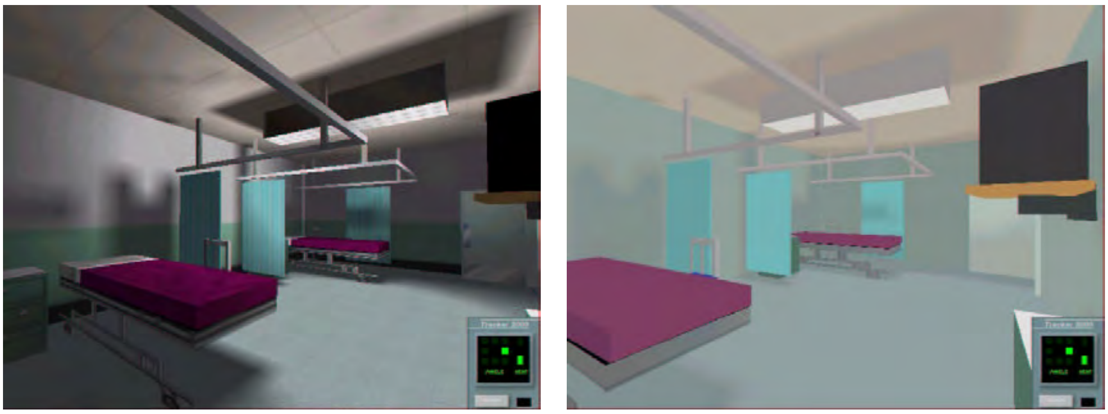


Figure C-8: Ward room (level 3)

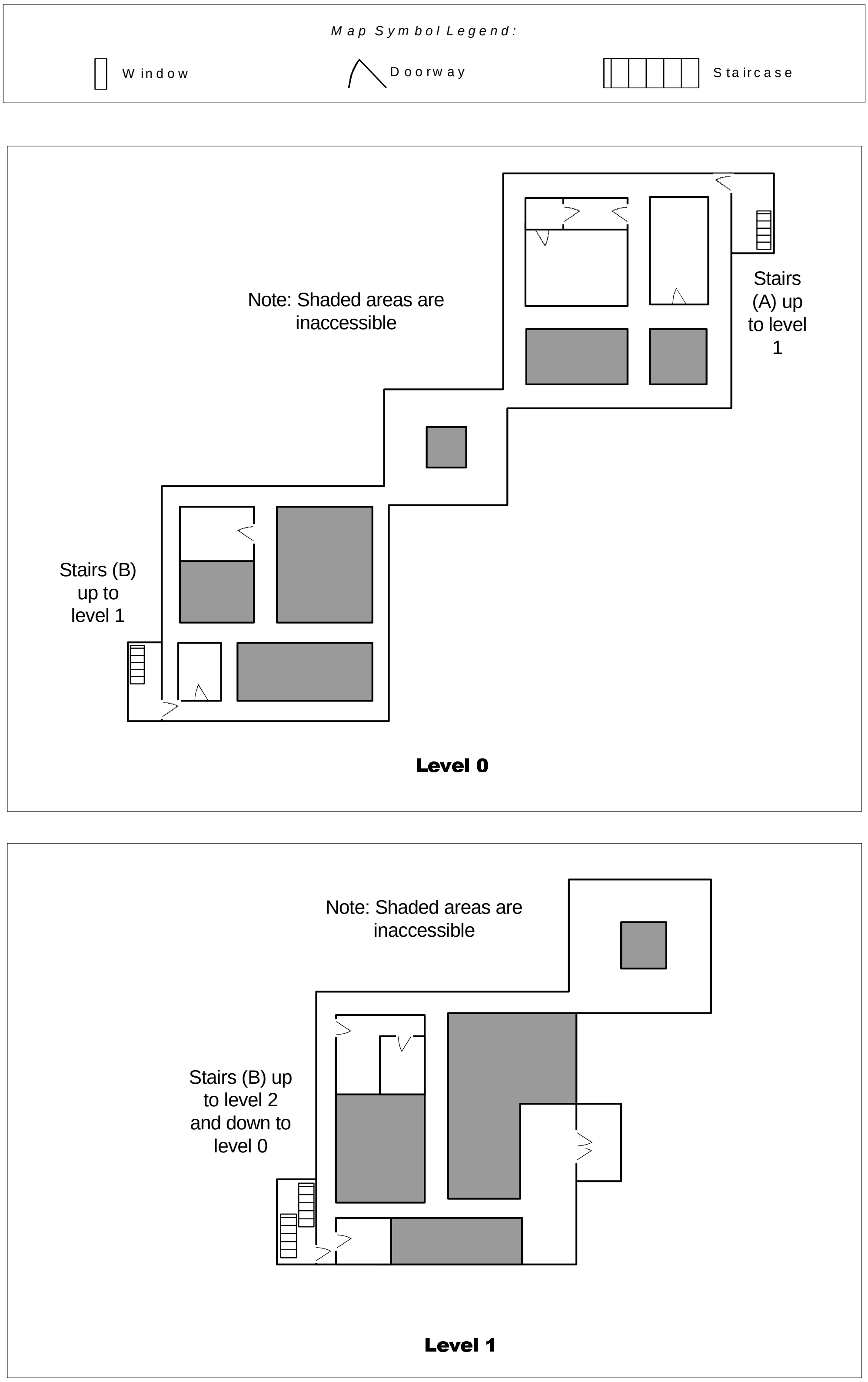


Figure C-9: Floorplans for the Hospital VE, Levels 0 and 1

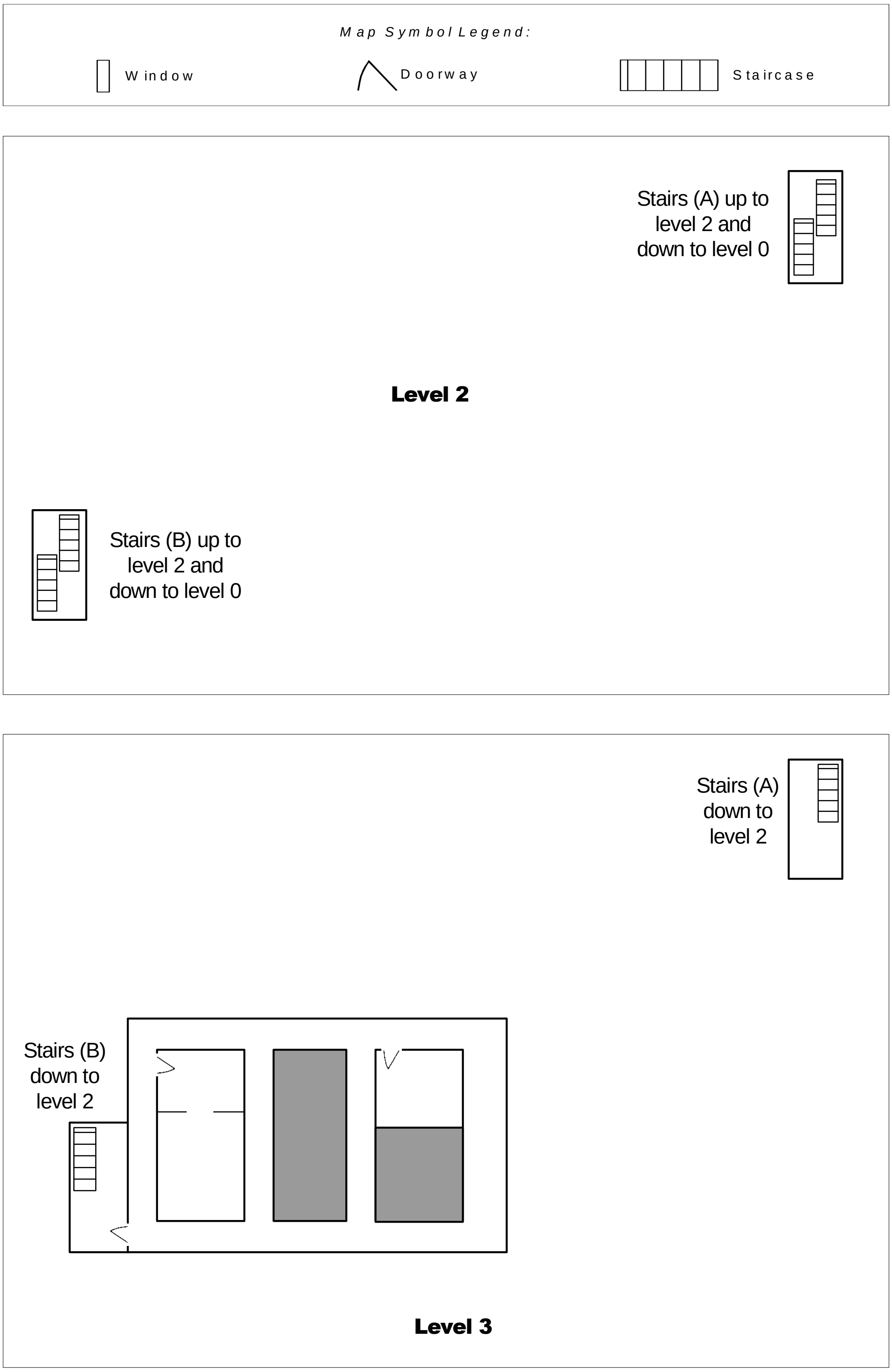


Figure C-10: Floorplans for the Hospital VE, Levels 2 and 3

Appendix D

Images from the Monastery VE

This appendix shows eight images taken from the high and low stimulus quality versions of the Monastery VE used in the experiments described in this dissertation. The VE is not based on any real place. In each case, the left-hand image is the high stimulus quality version, and the right-hand image is from the low stimulus quality version. Apart from the graphical differences, the high stimulus quality version included both positional and ambient sounds, while the low stimulus quality version did not include any sound.

Figure D-1: The entrance and stairs to first floor (level 1)

Figure D-2: Dining hall (level 1)

Figure D-3: Courtyard and chapel (level 1)

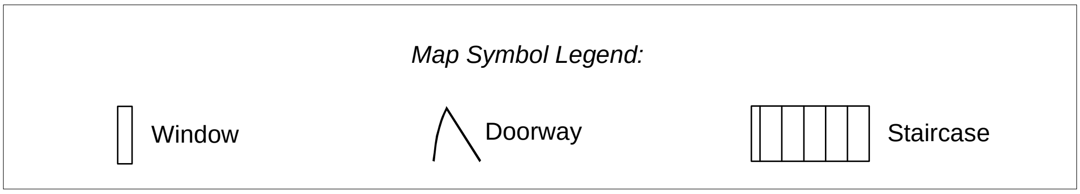
Figure D-4: Inside the chapel (level 1)

Figure D-5: Annexure to the chapel (level 1)

Figure D-6: Library (level 2)

Figure D-7: One of the bedrooms (first floor, level 2)

Figure D-8: Carpentry workshop (level 1)



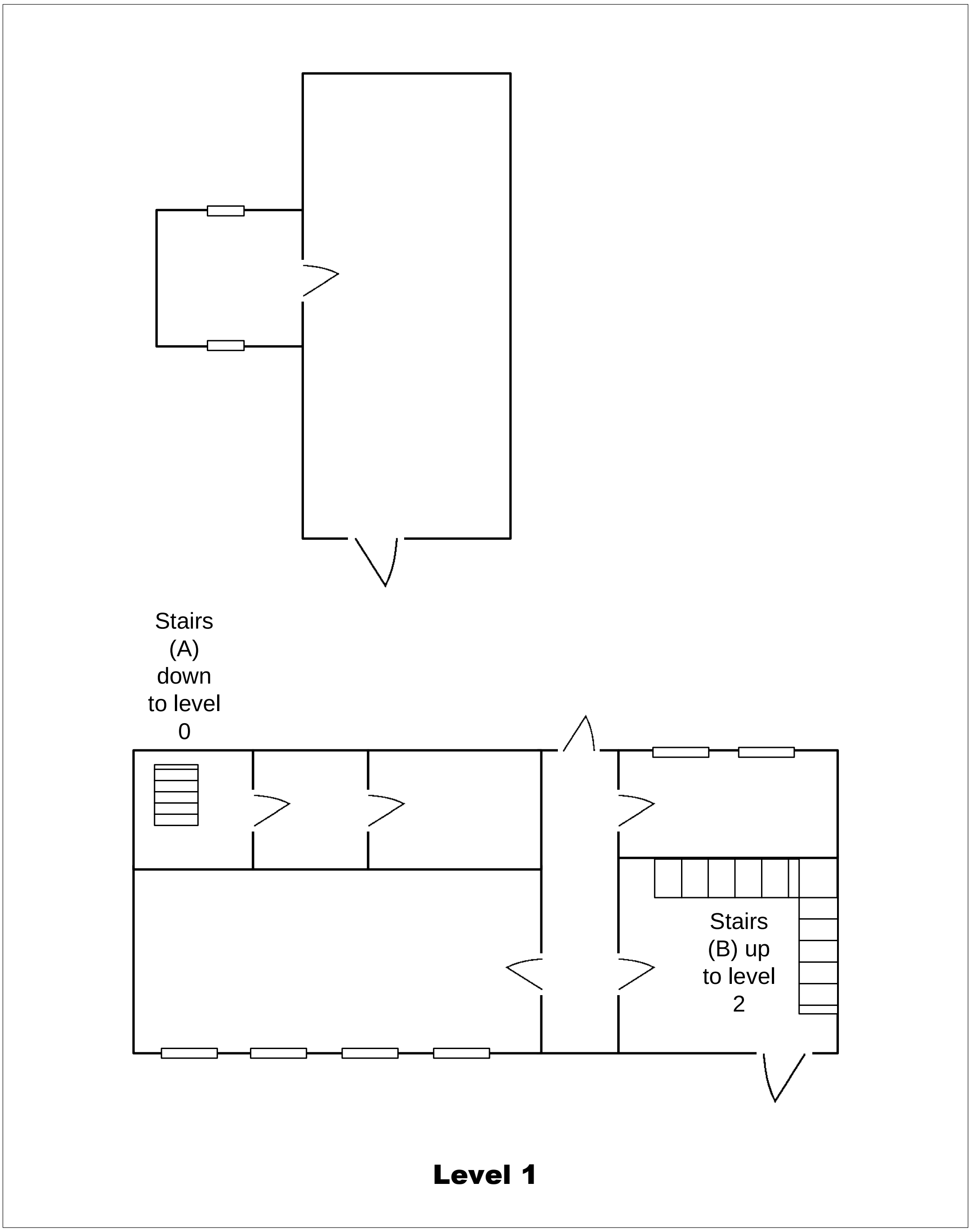


Figure D-9: Floorplans for the Monastery VE, Level 1

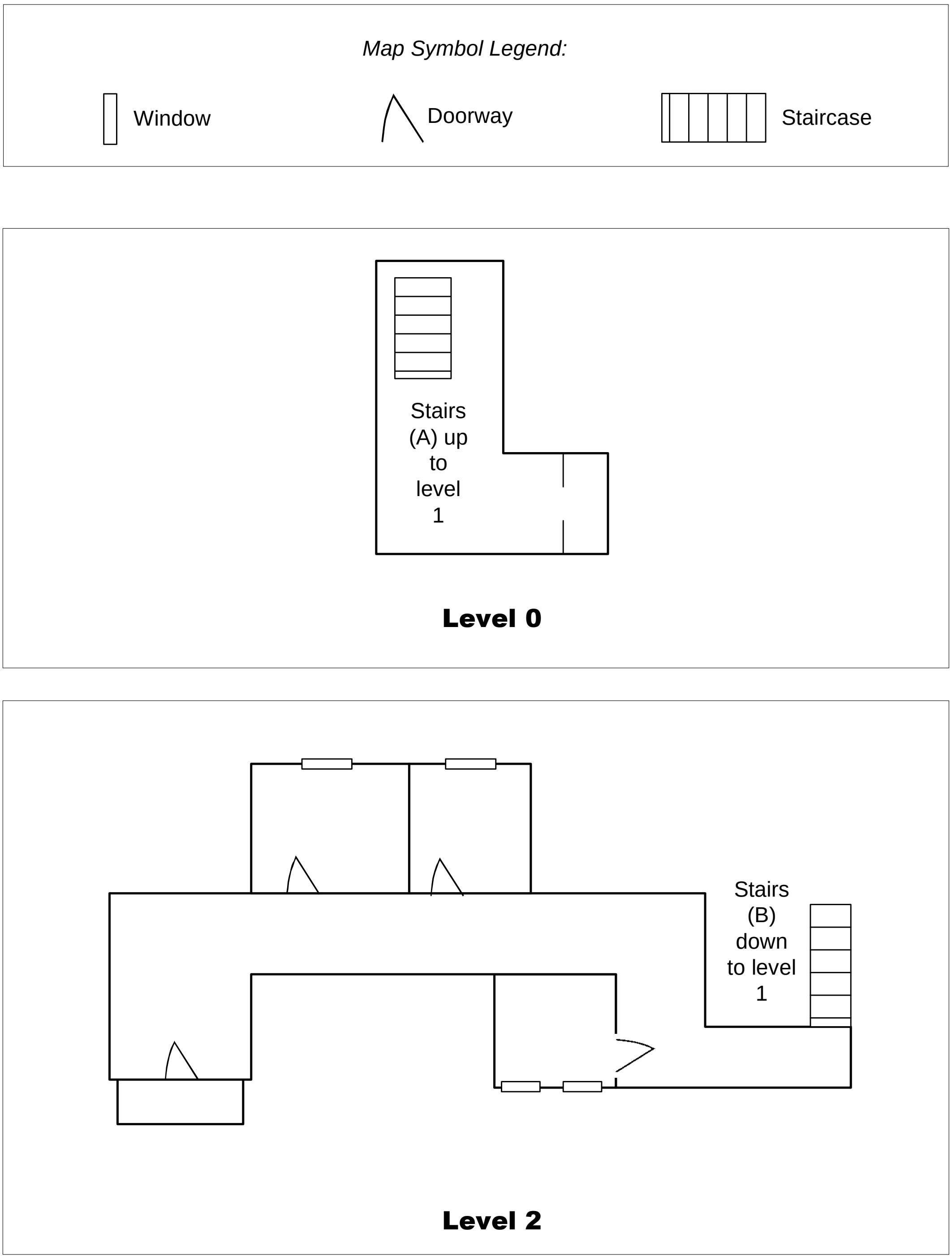


Figure D-10: Floorplans for the Monastery VE, Levels 0 and 2

Appendix E

Images from the Training VE

This appendix shows four images and floorplans from the training VE in which each subject learnt the controls and conventions of the DAVE engine. The training VE was designed not to fit any particular theme, but simply to act as a space in which subjects could learn to use the DAVE system.

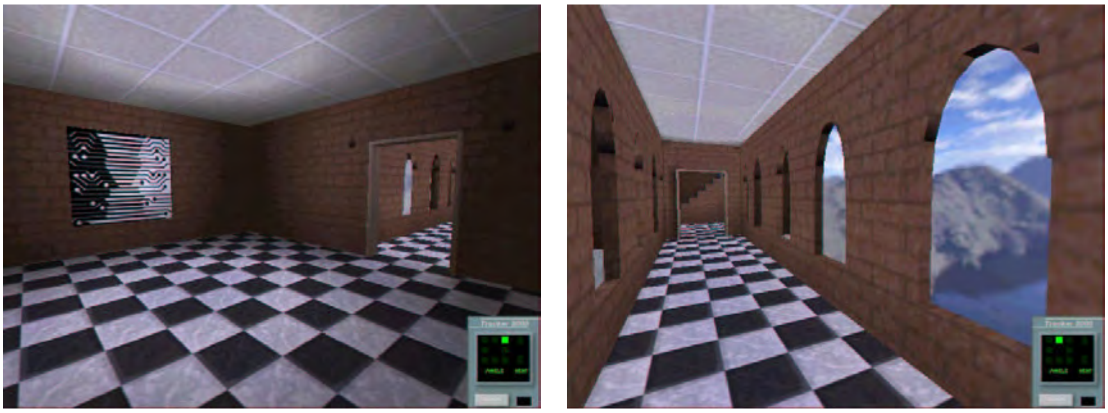


Figure E-1: Images from the first floor

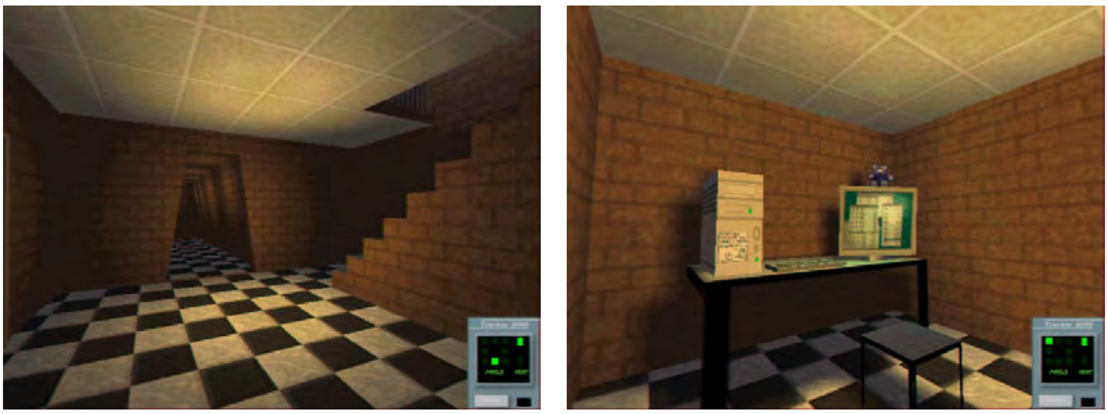


Figure E-2: Images from the basement

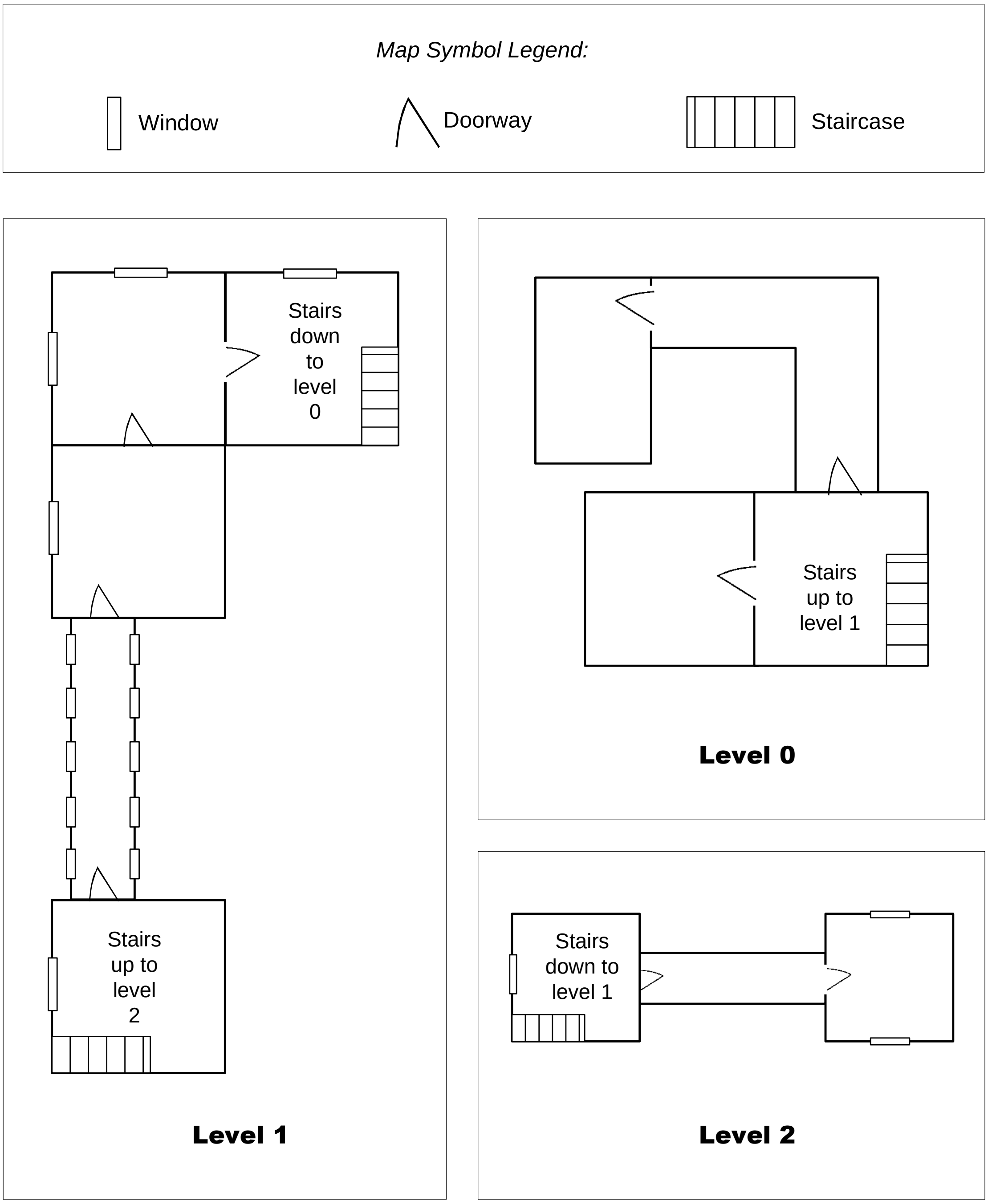


Figure E-3: Floorplans for the training level

Appendix F

Lists of COCI items

This appendix lists the COCI items used for each of the three environments (hospital, monastery and training). The monastery and hospital environments had COCI lists of ten items. The training environment only had six COCI items in its list. In each item, the corresponding word fragment is shown in italics, and the item keyed to the environment setting is displayed in bold.

F.1 Hospital COCI items:

1 *BL___D*

BLEND
BLOND
BLIND
BLOOD

2 *N ____E*

NETTLE
NOODLE
NEEDLE
NIBBLE

3 *PA____ ER*

PARAMETER
PAINKILLER
PATHFINDER
PAWNBROKER

4 *A____E*

AFFLUENCE
ALLOWENCE
AMBULANCE
APPLIANCE

6 *O_____ION*

OCCUPATION
OPPOSITION
OBLIGATION
OPERATION

7 *A______IC*

ANTISEPTIC
ANTISTATIC
AERONAUTIC
ALPHABETIC

8 *D_____E*

DEBATE
DIVORCE
DECLINE
DISEASE

9 *O_____N*

OCEAN
ONION
OLDEN
ORGAN

5 *D*____*R*

DOCTOR
DANCER
DEALER
DIGGER

10 *A*____*N*

ARTISAN
ASPIRIN
AILERON
AUCTION

F.2 Monastery COCI items

1 *B*____*E*

BIBLE
BUGLE
BAFFLE
BEAGLE

2 *C*____*LE*

CATTLE
CANDLE
CABLE
CYCLE

3 *W*___*D*

WORLD
WARD
WORD
WOOD

4 *M*____*K*

MONK
MARK
MILK
MASK

5 *CH*___*L*

CHISEL
CHANNEL
CHARNEL
CHAPEL

6 PR___T

PRODUCT
PROFIT
PRIEST
PRESENT

7 *A*____*Y*

ALLEY
AIRWAY
ARTERY
ABBEY

8 *C*____*T*

CHANT
CHART
COUNT
COURT

9 *S*____*E*

SCOPE
STONE
SLOPE
SCONE

10 D___T

DIET
DART
DUST
DUCT

F.3 Training COCI items

Note that there was no environment keyed items defined for the training world. The COCI in this world was used only for training, and the data collected was not analyzed.

1 *F____D*

FAKED
FADED
FIELD
FINED

2 *B____T*

BEST
BOAST
BLOT
BERET

3 *S____K*

SKULK
STICK
SLACK
SHIRK

4 *F____T*

FIGHT
FRONT
FACET
FRUIT

5 *T____R*

TEAR
TSAR
TOUR
TIER

6 *G____P*

GETUP
GRASP
GROUP
GRUMP

Appendix G

Questionnaires used

This appendix contains one copy of each of the questionnaires used in the two experiments. For details of which were used for what purpose, please refer to the chapter detailing the appropriate experiment.

G.1 Slater, Usoh & Steed questionnaire (SUS)

For the following questions, please circle the number which best represents your experience.

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

I had a sense of "being there" in the virtual environment:

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

2. To what extent were there times during the experience when the virtual environment was the reality for you?

There were times during the experience when the virtual environment was the reality for me…

1	2	3	4	5	6	7
At no time						Almost all the time

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

The virtual environment seems to me to be more like…

1	2	3	4	5	6	7
Images that I saw						Somewhere that I visited

4. During the time of the experience, which was the strongest on the whole, your sense of being in the virtual environment or of being elsewhere?

I had a stronger sense of…

1	2	3	4	5	6	7
Being elsewhere						Being in the virtual environment

5. Consider your memory of being in the virtual environment. How similar in terms of the *structure of the memory* is this to the structure of the memory of other *places* you have been today? By ‘structure of the memory’ consider things like the extent to which you have a visual memory of the virtual environment, whether that memory is in colour, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic in your imagination, and other such *structural* elements.

I think of the virtual environment as a place in a way similar to other places that I’ve been today…

1	2	3	4	5	6	7
Not at all						Very much so

6. During the time of your experience, did you often think to yourself that you were actually in the virtual environment?

During the experience I often thought that I was really standing in the virtual environment…

1	2	3	4	5	6	7
Not very often						Very much so

G.2 Presence Questionnaire (PQ)

Characterize your experience in the virtual environment by circling the appropriate number on the seven point scale, in accordance with the question and descriptive labels. Please consider the entire scale when making your responses, as the intermediate levels may apply. Answer the questions independently in the order in which they appear. Do not skip questions or return to a previous question to change your answer.

WITH REGARD TO THE ENVIRONMENT YOU HAVE JUST EXPERIENCED

1. How much were you able to control events?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			COMPLETELY

2. How responsive was the environment to actions that you initiated (or performed)?

1	2	3	4	5	6	7
NOT RESPONSIVE			MODERATELY RESPONSIVE			COMPLETELY RESPONSIVE

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

1	2	3	4	5	6	7
EXTREMELY ARTIFICIAL			BORDERLINE			COMPLETELY NATURAL

4. How much did the visual aspects of the environment involve you?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			COMPLETELY

5. How much did the auditory aspects of the environment involve you?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			COMPLETELY

6. How natural was the mechanism which controlled movement through the environment?

1	2	3	4	5	6	7
EXTREMELY ARTIFICIAL			BORDERLINE			COMPLETELY NATURAL

7. How compelling was your sense of objects moving through space?

1	2	3	4	5	6	7
NOT AT ALL			MODERATELY COMPELLING			VERY COMPELLING

8. How much did your experiences in the virtual environment seem consistent with your real world experiences?

1	2	3	4	5	6	7
NOT CONSISTENT			MODERATELY CONSISTENT			VERY CONSISTENT

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

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			COMPLETELY

10. How completely were you able to actively survey or search the environment using vision?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			COMPLETELY

11. How well could you identify sounds?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			COMPLETELY

12. How well could you localize sounds?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			COMPLETELY

13. How well could you actively survey or search the environment using touch?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			COMPLETELY

14. How compelling was your sense of moving around inside the virtual environment?

1	2	3	4	5	6	7
NOT COMPELLING			MODERATELY COMPELLING			VERY COMPELLING

15. How closely were you able to examine objects?

1	2	3	4	5	6	7
NOT AT ALL			PRETTY CLOSELY			VERY CLOSELY

16. How well could you examine objects from multiple viewpoints?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			EXTENSIVELY

17. How well could you move or manipulate objects in the virtual environment?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			EXTENSIVELY

18. How involved were you in the virtual environment experience?

1	2	3	4	5	6	7
NOT INVOLVED			MILDLY INVOLVED			COMPELTELY ENGROSSED

19. How much delay did you experience between your actions and expected outcomes?

1	2	3	4	5	6	7
NO DELAYS			MODERATE DELAYS			LONG DELAYS

20. How quickly did you adjust to the virtual environment experience?

1	2	3	4	5	6	7
NOT AT ALL			SLOWLY			LESS THAN ONE MINUTE

21. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?

1	2	3	4	5	6	7
NOT PROFICIENT			REASONABLY PROFICIENT			VERY PROFICIENT

22. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?

1	2	3	4	5	6	7
NOT AT ALL			INTERFERED SOMEWHAT			PREVENTED TASK PERFORMANCE

23. How much did the control devices interfere with the performance of assigned tasks or with other activities?

1	2	3	4	5	6	7
NOT AT ALL			INTERFERED SOMEWHAT			INTERFERED GREATLY

24. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			COMPLETELY

25. How completely were your senses engaged in this experience?

1	2	3	4	5	6	7
NOT ENGAGED			MILDLY ENGAGED			COMPLETELY ENGAGED

26. To what extent did events occurring outside the virtual environment distract from your experience virtual environment?

1	2	3	4	5	6	7
NOT AT ALL			MODERATELY			VERY MUCH

27. Overall, how much did you focus on using the display and control devices instead of the virtual experience and experimental tasks?

1	2	3	4	5	6	7
NOT AT ALL			MODERATELY			VERY MUCH

28. Were you involved in the experimental task to the extent that you lost track of time?

1	2	3	4	5	6	7
NOT AT ALL			SOMEWHAT			COMPLETELY

29. 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?

1	2	3	4	5	6	7
IMPOSSIBLE			MODERATELY DIFFICULT			VERY EASY

30. Were there moments during the virtual environment experience when you felt completely focused on the task?

1	2	3	4	5	6	7
NONE			OCCASIONALLY			FREQUENTLY

31. How easily did you adjust to the control devices used to interact with the virtual environment?

1	2	3	4	5	6	7
DIFFICULT			MODERATE			EASY

32. Was the information provided though different senses in the virtual environment (eg. Vision, hearing, touch) consistent?

1	2	3	4	5	6	7
NOT CONSISTENT			SOMEWHAT CONSISTENT			VERY CONSISTENT

G.3 The Immersive Tendencies Questionnaire (ITQ)

Indicate your preferred answer by circling the appropriate number. Please consider the entire scale when making your responses, as the intermediate levels may apply. For example, if your response is once or twice, the number “2” should be marked. If your response is many times but not extremely often, then the number “6” should be marked.

1. Do you easily become deeply involved in movies or TV dramas?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

2. Do you ever become so involved in a television program or book that people have problems getting you attention?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

3. How mentally alert do you feel at the present time?

1	2	3	4	5	6	7
NOT ALERT			MODERATELY			FULLY ALERT

4. Do you ever become so involved in a movie that you are not aware of things happening around you?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

5. How frequently do you find yourself closely identifying with the character in a story line?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

6. Do you ever become so involved in a video game that it is as if you are inside the game rather than moving a joystick and watching the screen?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

8. How physically fit do you feel today?

1	2	3	4	5	6	7
NOT FIT			MODERATELY FIT			EXTREMELY FIT

9. How good are you at blocking out external distractions when you are involved in something?

1	2	3	4	5	6	7
NOT VERY GOOD			SOMEWHAT GOOD			VERY GOOD

10. When watching sports, do you ever become so involved in the game that you react as if you were one of the players?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

11. Do you ever become so involved in a daydream that you are not aware of things happening around you?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

12. Do you ever have dreams that are so real that you feel disoriented when you awake?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

13. When playing sports, do you become so involved in the game that you lose track of time?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

14. How well do you concentrate on enjoyable activities?

1	2	3	4	5	6	7
NOT AT ALL			MODERATELY WELL			VERY WELL

15. How often do you play arcade of video games? (OFTEN should be taken to mean every day or every two days, on average)

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

16. Have you ever gotten excited during a chase or fight scene on TV or in the movies?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

17. Have you ever gotten scared by something happening on a TV show or in a movie?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

18. Have you ever remained apprehensive or fearful long after watching a scary movie?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

19. Do you ever become so involved in doing something that you lose all track of time?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

21. Do you ever get involved in projects or tasks, to the exclusion of other activities?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

22. How easily can you switch attention from the activity in which you are currently involved to a new and completely different activity?

1	2	3	4	5	6	7
NOT SO EASILY			FAIRLY EASILY			QUITE EASILY

23. How often do you try out new restaurants or new foods when presented with the opportunity?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

24. How frequently do you volunteer to serve on committees, planning groups or other civic or social groups?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

25. How often do you try new things or seek out new experiences?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

26. Given the opportunity, would you travel to a country with a different culture and a different language?

1	2	3	4	5	6	7
NEVER			MAYBE			ABSOLUTELY

27. Do you go on carnival rides or participate in other leisure activities (horse back riding, bungee jumping, snow skiing, water sports) for the excitement and thrills they provide?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

28. How well do you concentrate on disagreeable tasks?

1	2	3	4	5	6	7
NOT AT ALL			MODERATELY WELL			VERY WELL

29. How often do you play computer games?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

30. How many different video, computer, or arcade games have you become reasonably good at playing?

NONE	ONE	TWO	THREE	FOUR	FIVE	MORE THAN FIVE

31. Have you ever felt completely caught up in an experience, aware of everything going on and completely open to it all?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			FREQUENTLY

32. Have you ever felt completely focused on something, so wrapped up in it that one activity that nothing could distract you?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			FREQUENTLY

33. How frequently do you get emotionally involved in (angry, sad, or happy) in news stories that you see, read or hear?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

34. Are you easily distracted when involved in an activity or working on a task?

1	2	3	4	5	6	7
NEVER			OCCASIONALLY			OFTEN

G.4 Form 100

▶In **no more than 10 words**, name/describe the virtual environment you just saw:

▶How much fun did you have while in this virtual environment? (circle the number below which best describes how much fun you had)

1	2	3	4	5	6	7	8	9	10
No fun at all									A lot of fun

▶What year would you say the virtual environment you just saw was set in?

▶Do you think the virtual environment you just saw was based on a real place? (Yes/No)

▶Did the virtual environment change your mood? If so, what mood did it cause you to feel?

▶What time of day was it in the virtual environment (morning/midday/afternoon/night)

▶What was the weather like in the virtual environment?

▶How many other people do you think there were in the virtual environment?

▶If you traveled northwards in the virtual environment for an hour, what do you think you would find there?

G.5 Form 200

▶In **no more than 10 words**, name/describe the virtual environment you just saw:

▶How much fun did you have while in this virtual environment? (circle the number below which best describes how much fun you had)

1	2	3	4	5	6	7	8	9	10
No fun at all									A lot of fun

▶What year would you say the virtual environment you just saw was set in?

▶Do you think the virtual environment you just saw was based on a real place? (Yes/No)

▶Did the virtual environment change your mood? If so, what mood did it cause you to feel?

▶What time of day was it in the virtual environment (morning/midday/afternoon/night)

▶What was the weather like in the virtual environment?

▶How many other people do you think there were in the virtual environment?

▶If you traveled northwards in the virtual environment for an hour, what do you think you would find there?

▶Which of the two virtual environments you saw did you most enjoy? (The first one / The second one)

Appendix H

Priming materials

This appendix contains the priming materials. Three booklets were constructed: one keyed to the monastery, one keyed to the hospital, and one keyed to an unrelated theme. Although the following pages contain headers, footers and page numbers, the booklets given to the subjects did not.

H.1 Monastery-keyed priming booklet

Early monasteries originated in Egypt as places where wandering hermits gathered. These early "monks" lived alone,

but met in a common chapel. By the fifth century the monastic movement had spread to Ireland, where St. Patrick, the son of a Roman official, set out to convert the Irish to Christianity. The Irish monks spread Christianity into Cornwall, Wales, and Scotland. St. Ninian established a monastery at Whithorn in Scotland about 400 AD, and he was followed by St. Columba (Iona), and St. Aidan, who founded a monastery at Lindisfarne in Northumbria. 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.

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.

Although the details of daily life differed from one order to the next (as mentioned above), monastic life was generally one of hard physical work, scholarship and prayer. Some orders encouraged the presence of "lay brothers", monks who did most of the physical labour in the fields and workshops of the monastery so that the full-fledged monks could concentrate on prayer and learning.

Typically a monastery was in the charge of an abbot. The abbot was responsible for the souls of the monks, which often 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.

In addition to the church, a monastery had a number of other buildings. In addition to the dormitory and refectory, a monastery would often have an infirmary and a guest house. The buildings were often set around an area known as the cloister.

The day of a monk or nun, in theory at least, was regulated by regular prayer services in the abbey church. These services took place every three hours, day and night. When the services were over, monks would be occupied with all the tasks associated with maintaining a self-sustaining community.

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).

By 680 AD the Pope had sent a Benedictine monk from Saint Peter's in Rome to instruct monks how to chant prayers in the Benedictine style. This uniformity showed how patterns were changing. 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.

The abbey (the term for a monastery or nunnery) was under the authority of an abbot or abbess. The abbot could be a landless noble, who used the church as a means of social advancement. Under the abbot was the prior/prioress, who ran the monastery in the absence of the abbot, who might have to travel on church business. There could also be a sub-prior. Other officers included the cellerar (in charge of food storage and preparation), and specialists in the care of the sick, building, farming, masonry, and education.

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.

H.2 Hospital-keyed priming booklet

Understanding the E.R. Maze

The classic emergency room scene involves an ambulance screeching to a halt, a gurney hurtling through the hallway and 5 people frantically working to save a person's life with only seconds to spare. This *does* happen and it is not uncommon, but the majority of cases seen in a typical emergency department

aren't quite this dramatic. Let's look at a typical case to see how the normal flow of an emergency room works.

Imagine that it's 2AM, the kids and pets are asleep and you're dreaming about whatever it is that you dream about. Suddenly your ten year old daughter wakes you up to tell you that her belly hurts a lot worse than that little ache she had after dinner. This seems like something out of the ordinary, so you call your pediatrician. He tells you to go to your local hospital's Emergency Department. He is concerned about appendicitis because her pain is located in the right lower abdomen.

Triage

When you arrive at the Emergency Department, your first stop is Triage. This is the place where each patient's condition is prioritized, typically by a nurse, into three general categories. The categories are:

- Immediately life threatening
- Urgent but not immediately life threatening
- Less urgent

This categorization is necessary so that someone with a life threatening condition is not kept waiting because they arrived a few minutes later than someone with a more routine problem. The triage nurse records your daughter's vital signs (temperature, pulse, respiratory rate and blood pressure). She also gets a brief history of her current medical complaints, past medical problems, medications and allergies so that she can determine the appropriate triage category for her. Here you find out that your daughter's temperature is 101 degrees F.

Registration

Next stop is registration – not very exciting and rarely seen on TV! Here they obtain your daughter's vital statistics. You may also provide them with your insurance information, Medicare, Medicaid or your H.M.O. card. This step is necessary to develop a medical record so that your daughter's medical history, lab tests, X-rays, etc., will all be located on one chart that can be referenced at any time. The bill will also be generated from this information. Note that all patients must receive a medical screening exam regardless of their ability to pay.

If the patient's condition is life-threatening or if the patient arrives by ambulance, this step may be completed later at the bedside.

Examination Room

Now your daughter is brought back to the exam room. She promptly throws up in the bathroom. Perhaps this is more evidence for a diagnosis of appendicitis. She is now seen by an emergency department nurse who obtains more detailed information. The nurse gets her settled into a patient gown so that she can be examined properly and perhaps obtains a urine specimen at this time.

Some Emergency Departments have been subdivided into separate areas to better serve their patients. These separate areas can include - a pediatric ER, a chest pain ER, a fast track (for minor injuries and illnesses), trauma center (usually for severely injured patients) and an observation unit (for patients that do not require admission but require prolonged treatment or many diagnostic tests).

Once the nurse has finished her tasks, the next visitor is an Emergency Medicine physician. He gets a more detailed medical history about her present illness, past medical problems, family history, social history, and a complete review of all her body systems. He then formulates a list of possible causes of her symptoms. This list is called a **differential diagnosis**. The most likely diagnosis is then determined by the patient's symptoms and physical examination. If this is inadequate to determine the diagnosis, then diagnostic tests are required.

Diagnostic Tests

When the tricky diagnosis of appendicitis is considered, blood tests and a urinalysis are required. An I.V. line may be inserted at this time so that fluids can be given intravenously (through the vein) to replace fluids lost through vomiting. The patient's blood is put into different colored tubes, each with it own additive depending on the test being performed:

- A purple top tube is used for a complete blood count (CBC). A CBC measures 1) the adequacy of your red blood cells (to see if you are anemic), 2) The number and type of white blood cells (WBC's) (to determine the presence of infection), and 3) a platelet count (platelets area blood component necessary for clotting).
- A red top tube is used to test the serum (the liquid or non cellular half of your blood).
- A blue top tube is used to test your blood's clotting.

The tests in your daughter's case indicate that your daughter has an elevated WBC count. This is a sign of a bacterial infection, and bacterial infections are commonly associated with appendicitis.

At this point the emergency physician may request that your daughter not eat or drink anything . The reason is that appendicitis is treated by surgery and an empty stomach is desirable to prevent some complications of anesthesia.

Diagnosis and Treatment

When the emergency physician has all the information he can obtain, he makes a determination of the most likely diagnosis from his differential diagnosis. Alternately, he may decide that he does not have enough information to make a decision and may require more tests.

In this case he speaks to a general surgeon - the appropriate consultant in this case. The surgeon comes to see your daughter and performs a thorough history, physical exam, and review of her lab data. She examines your daughter's symptoms: pain and tenderness in the right lower abdomen, vomiting, low grade fever and elevated WBC count. These symptoms all point to appendicitis. The treatment of appendicitis is removal of the appendix, or an appendectomy. The surgeon explains the procedure, including the risks and benefits. You then sign a consent form to document this and permit her to operate on your daughter. You're nervous, yet relieved because you know what the problem is, and that your daughter will be feeling better once her appendix is removed.

H.3 Neutral priming booklet

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!

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 steampowered 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.

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.

Eventually, the moment arrived: we had about 140 lb/in?pressure, and the engine could be moved, literally under its own steam. We backed up to the headshunt, picked up the smaller brake van, and we were ready for practical instruction. First, student no. 1 drove while no. 2 (that’s me) fired; then no. 2 drove while no. 3 fired, and so on.

The student who was learning to fire was introduced to John, who had been with the Bluebell Line for 16 years. He showed us how to pick up the deceptively small amount of coal on the shovel, make a full swing, and turn the shovel’s handle to flip the coal to the desired place on the grate. It didn't need much effort for throwing, but that flip! Sometimes the coals would stay together in a clump and land in one spot, even if it was not quite the right spot, but sometimes they would spray right across the full width of the grate. I ran through my repertoire of curses.

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.

References

1. Alison, R.S., Harris, L.R., Jenkin, M., Jasiobedzka, U., & Zacher, J.E. (2001). Tolerance of temporal delay in virtual environments. *IEEE International Conference on Virtual Reality, 3.*

2. Anastasi, A., & Urbina, S. (1996). *Psychological Testing.* New York, NY: Prentice Hall.

3. Barfield, W., Rosenberg, C., Han, S.H., & Furness, T. A. (1993). Spatial situational awareness as a function of frame of reference, virtual eyepoint elevation, and geometric field of view. Proceedings of the 1993 SID International Symposium, Digest of Technical Papers, 107-110.

4. Barfield, W., & Weghorst, S. (1993). The sense of presence within virtual environments: a conceptual model. In G. Salvendy and M. Smith (Eds.), *Human-Computer Interaction: Software and Hardware Interfaces*. Amsterdam, The Netherlands: Elsevier Science Publishers.

5. Barfield, W., & Hendrix, C. (1995a). Factors affecting presence and performance in virtual environments. In R. M. Satava, K. Morgan, H. B. Sieburg, R. Mattheus, & J. P. Christensen (Eds.), *Interactive Technology and the New Paradigm for Healthcare.* Amsterdam, The Netherlands: IOS Press.

6. Barfield, W. & Hendrix, C. (1995b). The effect of update rate on the sense of presence within virtual environments. *Virtual Reality: Research, Development and Applications, 1,* 3-16

7. Barfield, W., Zeltzer, D., Sheridan, T.B., & Slater, M. (1995). Presence and performance within virtual environments. In W. Barfield & T. Furness(Eds.), *Virtual Environments and Advanced Interface Design.* London: Oxford University Press.

8. Beurle, R.L. (1956). Properties of a mass of cells capable of regenerating pulses. *Philosophical Transactions of the Royal Society, B, 240,* 55-94.

9. Biocca, F. (1992). Will simulation sickness slow down the diffusion of virtual environments technology? *Presence: Teleoperators and Virtual Environments, 3,* 334-343

10. Blake, E. H., Casanueva, J., & Nunez, D. (2000). Presence as a means for understanding user behaviour in virtual environments. *South African Computer Journal, 26,* 247-252.

11. Bower, G.H., Black, J.B., & Turner, T.J. (1979). Scripts in memory for text. *Cognitive Psychology, 11,* 177-220.

12. Brown, R. (1986) *Social Psychology*. New York, NY: The Free Press

13. Bystrom, K. E., Barfield, W., & Hendrix, C. (1999). A conceptual model of the sense of presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 8,* 241-244.

14. Carver, C. S., & Scheier, M. F. (1998). *On the Self-Regulation of Behavior*. New York, NY: Cambridge University Press

15. Cassell, J., & Thorisson, K. R. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in conversational agents. *Applied Artificial Intelligence, 13,* 519-538

16. Cassell, J., Torres, O., & Prevost, S. (1999). Turn taking vs. discourse structure: How best to model multimodal conversation. In Y. Wilks (Ed.), *Machine Conversations.* The Hague, The Netherlands: Kluwer Press.

17. Caudill, M., & Butler, C. (1990). *Naturally Intelligent Systems*. Cambridge, MA.: MIT Press

18. Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press

19. Coleman, L., & Kay, P. (1981). Prototype semantics. *Language, 57,* 26-44.

20. Collins, A.M., & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 407-428.

21. Cook, T.D., & Campbell, D.T. (1979) Quasi-experimentation: Design and analysis issues for field settings. Chicago, IL.: Rand McNally

22. Craik, K. (1943). *The Nature of Explanation.* New York, NY: Cambridge University Press

23. Cronbah, L. J. (1990). *Essentials of Psychological Testing*. New York, NY: Addison Wesley

24. 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*, University of Paderborn, Germany.

25. Durlach, N., & Slater, M. (2000). Presence in shared virtual environments and virtual togetherness. *Presence: Teleoperators and Virtual Environments, 9,* 214-217.

26. Eysenk, M. W., & Keane, M.T. (1995). *Cognitive Psychology: A Student's Handbook*. Hove, U.K.: Psychology Press

27. Fencott, C. (1999) Presence and the content of virtual environments. *Paper presented at the 2nd International Workshop on Presence,* University of Essex, Colchester, UK.

28. Fodor, J.A. (1983). *The modularity of the mind.* Cambridge, MA: MIT Press

29. Freeman, J., Avons, S.E., Pearson, D.E., & IJsselsteijn, W.A. (1999). Effects of sensory information and prior experience on direct subjective ratings of presence, *Presence: Teleoperators and Virtual Environments, 8,* 1-13.

30. 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,* 241-244.

31. Galambos, J.A., & Rips, L.J. (1982). Memory for routines. *Journal of Verbal Learning and Verbal Behavior, 21,* 260-281

32. Gentner, D. (1981). Verb structures in memory for sentences: Evidence for componential representation. *Cognitive Psychology, 13,* 56-83

33. Gibson, J.J. (1979). *The Ecological Approach to Visual Perception.* Boston, MA: Houghton-Mifflin

34. Gregory, R.J. (1991). *Psychological Testing: History, Principles and Applications.* Boston, MA.: Allyn and Bacon

35. Grossberg, S (1980). How does the brain build a cognitive code? *Psychological Review, 87,* 1-51.

36. Hebb, D.O. (1949). *The organization of behavior.* New York, NY: Wiley

37. Held, R.M., & Durlach, N.I. (1992). Telepresence. *Presence: Teleoperators and virtual environments, 1,* 109-112.

38. Hendrix, C., & Barfield, W. (1995). Presence in virtual environments as a function of visual and auditory cues. *Proceedings of the Virtual Reality Annual International Symposium '95,* 74-82.

39. Howell, D.C. (2002). *Statistical Methods for Psychology.* Pacific Grove, CA: Duxbury

40. Hughes, J.A. (1994). *The Philosophy of Social Research.* London, UK: Longman Group Limited

41. IJsselsteijn, W.A., de Ridder, H., Hamberg, R., Bouwhuis, D., & Freeman, J. (1998). Perceived depth and the feeling of presence in 3DTV, *Displays, 18,* 207-214.

42. IJsselsteijn, W. A., de Ridder, H., Freeman, J., & Avons, S. E. (2000). Presence: Concept, determinants and measurement. *Proceedings of the SPIE, Human Vision and Electronic Imaging V,* 3959-76.

43. IJsselsteijn, W.A., Freeman, J., de Ridder, H., Avons, S.E., & Pearson, D.E. (2000). Towards an objective corroborative measure of presence: Postural responses to moving video. *Paper presented at the 3rd International Workshop on Presence,* Delft, The Netherlands.

44. Jacobson, D. (2002). On theorizing presence. *Journal of Virtual Environments, 6*(1) Available online: http://www.brandeis.edu/pubs/jove/HTML/V6/presence.html

45. Johns, C.L., Nunez, D., Daya, M., Sellars, D., Casaneuva, J.S., & Blake, E.H. (2000) The interaction between individuals' immersive tendencies and the sensation of presence in a virtual environment. In J.D. Mulder & R. van Liere (Eds). *Virtual Environments 2000: Proceedings of the Eurographics Workshop.* London: Springer-Verlag.

46. 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.

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

48. Kohonen, T. (1989). Speech recognition based on topology-preserving neural maps. In I Aleksander (Ed.). *Neural Computing Architectures: The Design of Brain-like Machines.* Cambridge, MA: MIT Press.

49. Koronoski, J. (1967). *Integrative activity of the brain.* Chicago, IL: University of Chicago Press

50. Lessiter, J., Freeman, J., & Davidoff, J.B (2001). Really hear - The effects of audio quality on presence. *Paper presented at the 4th International Workshop on Presence,* Temple University, Philadelphia PA

51. Loehlin, J.C. (1998). Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis. Hillsdale, NJ: Lawrence Erlbaum Associates.

52. Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3.*

53. Lombard, M., & Ditton, T. (2000). Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument. *Paper presented at the 3rd International Workshop on Presence,* Delft, The Netherlands.

54. Loomis, J.M. (1992). Distal attribution and presence. *Presence: Teleoperators and Virtual Envronments, 1,* 113-119.

55. Mania, K. (2001). Connections between lighting impressions and presence in real and virtual environments. *Proceedings ACM AFRIGRAPH,* 119-125.

56. Mania, K. & Chalmers, A. (2001). The effects of levels of immersion on presence and memory in virtual environments: A reality centered Approach. *Cyberpsychology & Behavior, 4,* 247-264.

57. Mantovani, G., & Riva, G. (1999). “Real” presence: How different ontologies generate different criteria for presence, telepresence and virtual presence. *Presence: Teleoperators & Virtual Environments, 8,* 540-555

58. Martindale, C. (1981). *Cognition and consciousness*. Pacific Grove, CA.: Brooks/Cole

59. Martindale, C. (1990). *Cognitive Psychology: a Neural Network Approach*. Pacific Grove, CA.: Brooks/Cole

60. MCClelland, J. L., & Rumelhart, D.E. (1986) *Parallel distributed processing, Vol 2.* Cambridge, MA: MIT Press

61. McClelland, J. L., & Rumelhart, D. E. (1988). *Explorations in parallel distributed processing: A handbook of models, programs, and exercises.* Cambridge, MA: MIT Press.

62. Meehan, M., Insko, B., Whitton, M. & Brooks, F.P. (2002) Physiological Measures of Presence in Stressful Virtual Environments *Proceedings, SIGGRAPH 2002,* 645-652

63. Miller, G.A. (1956). The magic number seven, plus of minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81-93.

64. Miller, J. C., Sharkey, T. J., Graham, G. A., & McCauley, M. E. (1993). Autonomic physiological data associated with simulator discomfort. *Aviation, Space, and Environmental Medicine, 64,* 813-819

65. Nass C., Steuer J., & Tauber E. (1994), Computers are social actors. *Paper presented at ACM SIGCHI '94,* Boston, MA.

66. Neely, J.H. (1977). Semantic priming and retrieval from lexical memory: the roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General, 106,* 226-54

67. Neiser, U. (1967). *Cognitive Psychology.* New York: Appleton-Century-Crofts.

68. Nisbett, R.E. & Wilson, T.D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review, 84,* 231-259.

69. Norman, D. A. (1988). *The Psychology of Everyday Things.* New York, NY: Basic Books.

70. Nunez, D., & Blake, E.H. (2001). Cognitive presence as a unified concept of virtual reality effectiveness. *Proceedings ACM AFRIGRAPH,* 115-118

71. Nunnally, J. (1967). *Psychometric Theory.* New York, NY.: McGraw-Hill

72. Paivio, A. (1975). Nonmentalism. *Canadian Journal of Psychology, 29*, 263-291.

73. Pausch, R., Crea, T., & Conway, M. (1992). A literature survey for virtual environments: Military flight simulator visual systems and simulator sickness. *Presence: Teleoperators & Virtual Environments, 3,* 344-363

74. Pinker, S. (1995). *The Language Instinct.* New York, NY: Harper Perennial

75. Pollack, I., & Pickett, J.M. (1964). Intelligibility of excerpts from fluent speech: Auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior, 3,* 79-84.

76. Posner, M.I., & Snyder, C.R.R. (1975). Attention and cognitive control. In R L Solso (Ed) *Information processing and cognition.* Hillsdale, NJ: Erlbaum.

77. Prothero, J.D., Parker, D.E., Furness, T.A., & Wells, M.J. (1995). Towards a robust, quantitative measure for presence. *Proceedings of the Conference on Experimental Analysis and Measurement of Situation Awareness,* 359-366.

78. Regenbrecht, H., & Schubert, T. (2002). Real and illusory interactions enhance presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 11,* 425-434.

79. Reicher, G.M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 274-280

80. Rumelhart, D.E. (1975). Notes on a schema for stories. In D.G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York, NY: Academic Press

81. 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.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

82. Rumelhart, D.E., & Norman, D.A. (1981). Analogical processes in learning. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

83. Rumelhart, D.E., Hinton, G., & McClelland, J.L. (1986). A general framework for parallel distributed processing. In D. Rumelhart & J.L. McClelland (Eds), *Parallel Distributed Processing, Vol. 1: Explorations in the microstructure of cognition.* Cambridge, MA.: MIT Press

84. Salln?s, E. L. (1999). Presence in multimodal interfaces. *Paper presented at the 2nd International Workshop on Presence,* University of Essex, Colchester, UK.

85. 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 4th International Workshop on Presence,* Temple University, Philadelphia, PA.

86. Schank, R. & Abelson, R. (1977). *Scripts, plans, goals and understanding.* Hillsdale, NJ.: Erlbaum

87. Schloerb, D.W. (1995). A quantitative measure of telepresence. *Presence: Teleoperators and Virtual Environments, 4,* 64-80.

88. Schubert, T.W., Friedmann, F., & Regenbrecht, H.T. (1999). Embodied presence in virtual environments. In R. Paton & I. Neilson (Eds.), *Visual Representations and Interpretations,* London: Springer-Verlag.

89. Schubert, T.W., Friedmann, F., & Regenbrecht, H.T. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments, 10,* 266-281.

90. Schuemie, M.J., & van der Mast, C.A.P.G (1999). Presence: Interacting in virtual reality? In A. Nijholt, O. Donk, & B. van Dijk (Eds.) *Proceedings ,Twente Workshop on Language Technology 15,* 213-217.

91. 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,* 183-201.

92. Sheridan, T.B., Kruser, D.S., & S Deutsch. (1987). *Human Factors in Automated and Robotic Space Systems: Proceedings of a Symposium,* Symposium of the NRC Committee on Human Factors.

93. Sheridan, T.B. (1992). Musings of telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments, 1,* 120-125.

94. Sheridan, T.B. (1996). Further musings on the psychophysics of presence. *Presence: Teleoperators and Virtual Environments, 5,* 241-246

95. Sheridan, T.B. (1999). Descartes, Heidegger, Gibson and God: Towards an eclectic ontology of presence. *Presence: Teleoperators and virtual environments, 8,* 551-559.

96. Singer, M.J. & Witmer, R.G. (1999). On selecting the right yardstick. *Presence: Teleoperators and Virtual Environments, 8,* 566-573

97. Slater, M., & Usoh, M. (1993). Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators and virtual environments, 2,* 221-233

98. Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 3,* 130-144

99. Slater, M., Usoh, M., & Chrysanthou, Y. (1995). The influence of dynamic shadows on presence in immersive virtual environments. In M. G?bel (Ed.), *Virtual Environments '95: Selected Papers of the Eurographics Workshops.* New York, NY: SpringerWien.

100. Slater, M., Usoh, M. & Steed, A. (1995). Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interface, Special Issue on Virtual Reality Software and Technology, 2,* 201-219

101. Slater, M., & Wilbur, S. (1995). Through the looking glass world of presence: A framework for immersive virtual environments. In M. Slater (Ed.), *FIVE 95 Framework for Immersive Virtual Environments, 2*. QMW, University of London.

102. Slater, M., Usoh, M., Benford, S., Snowdon, D., Brown, C., Rodden, T., Smith, G. & Wilbur, S. (1996). Distributed Extensible Virtual Reality Laboratory (DEVRL). In M. Goebel, P. Slavik, & J.J. van Wijk (Eds). *Virtual Environments and Scientific Visualisation '96*. London: Springer Computer Science

103. Slater, M. (1999). Measuring presence: a response to the Witmer and Singer presence questionnaire. *Presence: Teleoperators and Virtual Environments, 8,* 560-565.

104. Slater, M., & Steed, A. (2000). A virtual presence counter. *Presence: Teleoperators and Virtual Environments, 9,* 413-434.

105. Slater, M., Sadagic, A., Usoh, M., & Shroeder, R. (2000) Small group behaviour in virtual and real environments: a comparative study. *Presence: Teleoperators and Virtual Environments, 9,* 37-51

106. Slater, M. (2002a). Course 49: Understanding Virtual Environments: Immersion, Presence, and Performance. *Siggraph 2002 Course Notes*

107. Slater, M. (2002b). Presence and the sixth sense. *Presence: Teleoperators and Virtual Environments,11,* 435-439.

108. Smith, S., Marsh, T., Duke, D., & Wright, P. (1998). Drowning in immersion. *Paper presented at UK-VRSIG'98,* Exeter, U.K.

109. Stark, L. W. (1996). The effects of pictorial realism, delay of visual feedback, and observer interactivity an the subjective sense of presence. *Presence: Teleoperators and Virtual Environments, 5,* 263-273.

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

111. Teuscher, C. (2001). *Turing's connectionism: an investigation of neural network architectures*. Heidelberg, Germany: Springer

112. Thie, S., & Wijk, J. (1998). A General Theory on Presence: Experimental Evaluation of Social Virtual Presence in a Decision Making Task. *Paper presented at Presence in Shared Virtual Environments Workshop*, University College London, London.

113. Thorndyke, P.W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology, 9,* 77-110.

114. Towell, J., & Towell, E. (1997). Presence in text-based networked virtual environments or "MUDS". *Presence: Teleoperators and Virtual Environments, 6,* 590-595

115. Triesman, A.M. (1964). Verbal cues, language and meaning in selective attention. *American Journal of Psychology, 77,* 206-219

116. Usoh, M., Arthur, K., Whitton, M.C., Bastos, R., Steed, A., Slater, M. & Brooks, F.P. (1999). Walking > walking-in-place > flying, in virtual environments. *Proceedings, SIGGRAPH99,* 359-364

117. Waller, D., Hunt, E. & Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence: Teleoperators and Virtual Environments, 7,* 129 – 143.

118. Williams, A. & Weisstein, N. (1978). Line-segments are perceived better in a coherent context than alone: An object-line effect in visual perception. *Memory and Cognition, 6,* 85-90.

119. Winn, W. (1993). A conceptual basis for educational applications of virtual reality. *HITL Technical Report R-93-9.* Human Interface Technology Laboratory, University of Washington.

120. Witmer, B.G., & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7,* 225-240

121. Zahoric, P. & Jenison, R.I. (1998). Presence as being-in-the-world. *Presence: Teleoperators and virtual environments, 7,* 78-89.

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