# Depth of Presence in Virtual Environments

**Article** *in* Presence Teleoperators & Virtual Environments · January 1994

DOI: 10.1162/pres.1994.3.2.130 · Source: DBLP



Mel Slater
University of Barcelona


Anthony Steed
University College London

# Depth of Presence in Virtual Environments

Mel Slater, Martin Usoh, Anthony Steed,

Department of Computer Science, and

London Parallel Applications Centre,

QMW, University of London,

Mile End Road,

London E1 4NS, UK.

## Abstract

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

environments by using the virtual HMD, and a negative association for those who were transported through doors. Finally, four of the subjects moved their real left arm to match movement of the left arm of the virtual body displayed by the system. These four scored significantly higher on the kinesthetic representation system than the remainder of the subjects.

**Keywords**

Presence, telepresence, virtual reality, virtual environments, neuro-linguistic programming (NLP).

## 1. Introduction

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

Sensors placed on the body of human participants map real body movements onto corresponding movements of their self-representation in the virtual world. We call this self-representation a virtual body (VB). We call a computer system that supports such experience an "immersive virtual environment" (IVE). It is immersive since it immerses a representation of the person's body (the VB) in the computer generated environment. Immersion can lead to *presence*, the participant's sense of "being there" in the virtual environment. The psychological sense of presence may be considered as an emergent property of an IVE. It is important to understand the factors that contribute to this, and a means of quantifying the concept of presence itself. It has been argued (Steuer, 1992) that the very definition of virtual reality involves presence: "A virtual reality is defined as a real or simulated environment in which a perceiver experiences telepresence"[1].

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

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

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

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

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

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

## 2. Indicators of Presence

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

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

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

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

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

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

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

### 3. Stacking Environments

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

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

Now IVEs provide the possibility of simulation. Consider that while in $T(R)$ the subject repeats a simulation of the procedures (for example, donning a HMD) that are equivalent to T. (Of course, these cannot be identically simulated). We denote the resulting environment by $T(T(R))$ ? ?$T^2(R)$ . This procedure can be repeated over and over - leading to the idea of environment $T^i(R)$ ? ?$E_i$ (say) , with ?$E_0 = R$. We call this the process of *stacking environments*, and $E_i$ is at *depth* i in the stack.

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

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

so that

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

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

## 4. Experimental Design and Procedures

### 4.1 Factors Expected to Influence Presence

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

**(a) Laws of Physics**

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

**(b) Visual Cliff**

The earlier experiment also suggested strongly that presence was increased when subjects were presented to a scenario where they were standing on a narrow ledge over a precipice, reminiscent of E.J. Gibson's visual cliff experiment (Gibson and

Walk, 1960). Subjects exhibited strong physical reactions (legs shaking, exclamations), and later reported this aspect of their experience as increasing their sense of presence.

**(c) Virtual Actors Responding to the Subject**

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

**(d) Subjective Factors**

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

## 4.2 Factorial Design

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

**Table 1**

Experimental Design

**(a) Stacking Environments Using HMD**

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

**(b) Going through Environments via Doors**

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

**Stacking Environments**

The entire scenario consisted of up to six different environments. Subjects could pass from environment to environment according to only one of two methods, depending on whether they had been assigned to Table 1(a) or Table 1(b). Those (12) assigned to

Table 1(a) would see a representation of a HMD while in a VE, and would learn to transport from one level to the next by donning the virtual HMD. Those assigned to 1(b) would simply go through a door to the next environment.

**Gravity**

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

**Virtual Actor**

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

**Visual Cliff**

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

**Stacking Level**

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

### 4.3 Story Line

From the experience with our earlier pilot studies there was a strong requirement to minimise contact between experimenters and subjects once the experiment had started.

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

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

### 4.4 Procedures

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

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

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

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

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

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

### 4.5 Equipment and Representation

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

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

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

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

### 4.6 Reinforcement of Environment Transitions

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

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

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

## 5. Instruments

### 5.1 Representation Systems and Perceptual Position

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

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

| The way I would think of my friend is... | Rank each answer:<br>1= most likely,<br>2 = next most likely,<br>3 = least likely |
|---|---|
| I mentally hear the voice and laughter of my friend. | |

| | |
|---|---|
| I feel as if my friend's presence is close to me. | |
| I see my friend's appearance. | |

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

Similarly, perceptual position was considered by questions such as:

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

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

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

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

### 5.2 Presence

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

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

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

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

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

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

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

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

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

## 6. Results

### 6.1 Summary

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

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

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

(c) No other factors were statistically significant.

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

### 6.2 Logistic Regression

**Dependent Variable**

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

**Independent Variables**

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

**Explanatory Variables**

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

This situation may be treated by logistic regression (Cox, 1970; Baker et. al., 1986), where the dependent variable is binomially distributed, with expected value related by the logistic function to a linear predictor.

Let the independent and explanatory variables be denoted by $x_1, x_2, \ldots, x_k$. Then the linear predictor is an expression of the form:

$$?_i ??? ?_0 ?+ \sum_{j=1}^{k} ?_j x_{ij} \quad (i = 1,2,\ldots,N)?$$

where N (=23) is the number of observations. The logistic regression model links the expected value $E(p_i)$ to the linear predictor as:

$$E(p_i) = \frac{n}{1+\exp(-?_i)}$$

where n (=3) is the number of binomial trials per observation.

Maximum likelihood estimation is used to obtain estimates of the ? coefficients. The deviance (minus twice the log-likelihood ratio of two models) may be used as goodness of fit significance test, comparing the null model ($?_j$ ?= 0, j = 1,...k) with any given model. The change in deviance for adding or deleting groups of variables may also be used to test for their significance. The (change in) deviance has an approximate $?^2$ ?distribution with degrees of freedom dependent on the number of parameters (added or deleted).

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

**Table 2**

Logistic Regression Equations

$\hat{?}$ = fitted values for the presence scale

(Coefficients are given to 1 d.p.)

V = Visual, A = Auditory, K = Kinesthetic, counts

L = number of levels

(a) Including Auditory

| **Group** | **Model** |
|---|---|
| No virtual HMD | $\hat{?}$ = 2.9 - 0.6*A - 0.7*L |
| With virtual HMD | $\hat{?}$ = -0.8 - 0.6*A + 0.3*L |

Overall Deviance = 21.916, d.f. = 18

$?^2$ ?at 5% on 18 d.f. = 28.869

| **Deletion of Model Term** | **Change in Deviance** | **Change in d.f.** | $?^2$ ?**at 5% level** |
|---|---|---|---|
| A | 8.06 | 1 | ?????? |
| HMD.L | 5.941 | 2 | ?????? |

(b) Including Visual and Kinesthetic

| **Group** | **Model** |
|---|---|
| No virtual HMD | $\hat{?}$ = -3.1 + 0.9*V + 0.5*K - 0.7*L |

| | |
|---|---|
| With virtual HMD | $\hat{?}$ = -7.5 + 0.9*V + 0.5*K + 0.4*L |

Deviance = 19.416, d.f. = 17
?$^2$ ?at 5% on 17 d.f. = 27.587

| **Deletion of Model Term** | **Change in Deviance** | **Change in d.f.** | **?$^2$ ?at 5% level** |
|---|---|---|---|
| V | 10.25 | 1 | ?????? |
| K | 4.067 | 1 | ?????? |
| HMD.L | 6.961 | 2 | ?????? |

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

**6.3 Discussion**

In the exploratory statistical analysis of the earlier pilot experiment, presence was measured by the responses to a question similar to 4 in Section 5.2. This was not satisfactory statistically, since as we stated at the time, an ordinal scale response was treated as a measured variable in the regression analysis. In the current paper we are on safer grounds, since our variable may be treated as the number of "successes" (6 or 7 scores) in three trials (the three questions of Section 5.2). However, the binomial model assumes independence between the trials, which is not obvious in the present context. The questions were each separated by two others in the questionnaire, and to a respondent, not knowing the purposes of the study, and not aware of the concept of presence, it would be reasonable to assume that answers did not directly influence one another, so that the "trials" were independent. (The correlations based on the full 24 sets of observations are 0.46 between questions 4 and 7, 0.62 between questions 4 and 10, and 0.50 between 7 and 10).

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

**Table 3**

Frequency Table for Dependent Variable

| **Count** | **Number of 1-7 scores = 4** | **Number of 1-7 scores = 5** |
|---|---|---|
| 0 | 7 | 12 |
| 1 | 4 | 5 |
| 2 | 6 | 5 |
| 3 | 6 | 1 |
| Total | 23 | 23 |

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

**Table 4**
Logistic Regression Equations
including VK = V? K

| Group | Model |
|---|---|
| No virtual HMD | $\hat{?}$ = 0.5 + 0.2*VK - 0.9*L |
| With virtual HMD | $\hat{?}$ = -4.0 + 0.2*VK + 0.3*L |

Overall Deviance = 22.510, d.f. = 18
$?^2$ ?at 5% on 18 d.f. = 28.869

| Deletion of Model Term | Change in Deviance | Change in d.f. | $?^2$ ?at 5% level |
|---|---|---|---|
| VK | 7.468 | 1 | ?????? |
| HMD.L | 9.09 | 2 | ?????? |

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

## 7. Presence and VK Association

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

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

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

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

**Table 5**
Mean and Standard Deviation of K score by left arm matching
$|t| = 2.756$. t on 21 d.f. = 2.080 at 5%

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

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

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

## 8. Conclusions

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

It has been suggested[2] that the reason for finding no significant associations between presence and the virtual actor or visual cliff independent variables, may be due to these being presented as solitary incidents. After leaving the VE subjects are required to make an evaluation of the entire experience, based only on these single experiences. This is important, and subsequent experimental designs should take this criticism into account. However, the visual cliff and virtual actor events always occurred at the end of the session, so that they would be more likely to be remembered and effective in influencing the final state of the subjects.

[2]This and the next paragraph are based on comments made by a referee who reviewed an earlier draft of this paper.

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

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

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

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

## Acknowledgements

We thank the anonymous referees for their detailed comments which have helped in production of the final version of this paper, and Allan Davison for helpful comments and suggestions. This work is funded by the U.K. Science and Engineering Research Council (SERC), and Department of Trade and Industry, through grant CTA/2 of the London Parallel Applications Centre. This is a collaborative project with Thorn Central Research Laboratories (CRL Ltd), and DIVISION Ltd. Anthony Steed is supported by a SERC research studentship.

## References

Baker, R.J. (1986) The GLIM System: User's Guide, Release 3.77, Royal Statistical Society, Numerical Algorithms Group (NAG).

Bandler, R. and J. Grinder (1979) Frogs into Princes, Real People Press.

Barfield, W. and S. Weghorst (1993) The Sense of Presence Within Virtual Environments: A Conceptual Framework, in Human-Computer Interaction:Software and Hardware Interfaces, Vol B, edited by G. Salvendy and M. Smith, Elsevier Publisher, 699-704, 1993.

Cox, D.R. (1970) Analysis of Binary Data, London: Menthuen.

Ellis, S.R. (1991) Nature and Origin of Virtual Environments: A Bibliographic Essay, Computing Systems in Engineering, 2(4), 321-347.

Fairchild, K.M., Beng Hai Lee, J. Loo, H. Ng, L. Serra (1993) The Heaven and Earth Virtual Reality: Designing Applications for Novice Users, IEEE Virtual Reality Annual International Symposium (VRAIS), September 18-22, Seattle, Washington, 47-53.

Gibson, E.J. and R.D. Walk (1960) The "visual cliff", Scientific American, 202, 64-71.

Heeter, C. (1992) Being There: The Subjective Experience of Presence, Presence: Teleoperators and Virtual Environments, 1(2), spring 1992, MIT Press, 262-271.

Held, R.M. and N.I. Durlach (1992) Telepresence, Presence: Teleoperators and Virtual Environments, 1, winter 1992, MIT Press, 109-112.

Loomis, J.M. (1992a) Distal Attribution and Presence, Presence: Teleoperators and Virtual Environments, 1, winter 1992, MIT Press, 113-119.

Loomis, J.M. (1992b) Presence and Distal Attribution: Phenomenology, determinants, and assessment, SPIE 1666 Human Vision, Visual Processing and Digital Display III, 590-594.

McGreevy, M.W. (1993) The Presence of Field Geologists in a Mars-Like Terrain, Presence: Teleoperators and Virtual Environments, 1(4) fall 1993, MIT Press, 375-403.

Sheridan, T.B. (1992) Musings on Telepresence and Virtual Presence, Presence: Teleoperators and Virtual Environments, 1, winter 1992, MIT Press, 120-126.

Slater, M., A. Steed and M. Usoh (1993) The Virtual Treadmill: A Naturalistic Metaphor for Navigation in Immersive Virtual Environments, Proceedings of the Eurographics Workshop on Virtual Reality, ed. by M. Goebel, Barcelona, September 1993.

Slater, M. and M. Usoh (1992) An Experimental Exploration of Presence in Virtual Environments, Department of Computer Science, QMW University of London.

Slater, M. and M. Usoh (1993) Presence in Immersive Virtual Environments, IEEE Virtual Reality Annual International Symposium (VRAIS), September 18-22, Seattle, Washington, 90-96.

Slater, M. and M. Usoh (1994) Representation Systems, Perceptual Position and Presence in Virtual Environments, Presence: Teleoperators and Virtual Environments, MIT Press, (in press).

Steuer, J. (1992) Defining Virtual Reality: Dimensions Determining Telepresence, Journal of Communication 42(4), 73-93.

## Appendix A

## Brief Description of the Scenes

**Scene 0**

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

**Scene 1**

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

**Scene 2**

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

**Scene 3**

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

**Scene 4**

A kitchen scene with cupboards and cooker

**Scene 5**

A bar scene with bar and bar furniture.

**Scene 6**

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

**Figure 1 (2 photos)**

The virtual HMD is being examined by a participant.