Immersion, Presence, and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess

Mel Slater, Vasilis Linakis[1], Martin Usoh, Rob Kooper[2]

Department of Computer Science
University College London,
Gower Street,
London WC1E 6BT
UK.

Abstract

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

Keywords

Virtual reality, virtual environments, immersion, presence, task performance, human factors, Tri-Dimensional Chess.

1. Introduction: Is VR better than a workstation?

This paper describes an experiment to compare task performance of twenty-four subjects with respect to immersive and non-immersive participation and interaction in a virtual environment (VE). Subjects participated in and observed a sequence of events played out in relation to a complex geometrical structure, and their task was the subsequent reproduction of those events with respect to the real world equivalent of that structure. Half the subjects participated in a visually immersive VE, and the remainder in a non-immersive VE. A secondary independent variable was the realism of the displayed environment. The task involved the reproduction of moves of pieces in the Tri-Dimensional chess game popularised in the TV series Star Trek. One inspiration of this work was the question posed by Mizell, Jones, Jackson and Picket (1995): "Is VR better than a workstation?" For our work this question breaks down into a number of components:

(a) Cognition of Geometric Structure

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

---

[1] Formerly Department of Computer Science, QMW University of London, where this work was carried out.

[2] Current address: College of Computing, Georgia Institute of Technology, USA.

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

(b) Knowledge Transfer

Here the question is whether skills or knowledge gained in a virtual environment can be successfully transferred to the real world. Suppose that a person learns to perform some task in a virtual environment, does immersion improve the chance of transfer of such knowledge to the real world? A previous attempt to study this (Slater, Alberto and Usoh, 1995) involved a number of subjects who walked through a virtual building immersively (with a head-slaved head-mounted display), and control group subjects who did the same non-immersively (looking at a monitor). The task involved finding a particular object within the virtual building. A response variable was the time it took them to find this object in the corresponding real building. No significant difference was found between performance for immersed and non-immersed subjects. However, the evidence suggested that those individuals with a high sense of presence, whether in an immersive or non-immersive system, achieved better performance overall. In this study also there was no interaction with objects in the virtual environment other than walkthrough.

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

(c) Immersion and Performance Within the Virtual Environment

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

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

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

2. Immersion, Presence and Task Performance

2.1 Immersion and Presence

In reports of earlier studies we have made a distinction between immersion and presence (Slater, Usoh and Steed, 1995). Immersion refers to what is, in principle, a quantifiable description of a technology. It includes the extent to which the computer displays are extensive, surrounding, inclusive, vivid and matching. The displays are more *extensive* the more sensory systems that they accommodate. They are *surrounding* to the extent that information can arrive at the person's sense organs from any (virtual) direction, and the participant can turn towards that direction receiving the appropriate directional sensory signals. The notion of surrounding also includes the greater the reproduction of the natural modes of sensory presentation (visual and auditory stereopsis for example). They are *inclusive* to the extent that all external sensory data (from physical reality) is shut out. Their *vividness* is a function of the variety and richness of the sensory information they can generate (Steuer, 1992). Vividness is concerned with the richness, information content, resolution and quality of the displays. Finally, immersion requires that there is *match* between the participant's proprioceptive feedback about body movements, and the information generated on the displays. A turn of the head should result in a corresponding change to the visual display, and, for example, to the auditory displays so that perceived sound direction is invariant to the orientation of the head. Matching requires body tracking, at least head tracking, but generally the greater the degree of body mapping, the greater the extent to which the movements of the body can be accurately reproduced.

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

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

Our general hypothesis is that presence is an increasing function of two orthogonal variables. The first variable is the extent of the match between the displayed sensory data and the internal representation systems and subjective world models typically employed by the participant. Although immersion is increased with the vividness of the displays, as discussed above, we must also take into account the extent to which the information displayed allows individuals to construct their own internal mental models of reality. For example, a vivid visual display system might afford some individuals a sense of "presence", but be unsuited for others in the absence of sound (Slater, Usoh and Steed, 1994). The second variable is the extent of the match between proprioception and sensory data. The changes to the display must ideally be consistent with and match through time, without lag, changes caused by the individual's movement and locomotion - whether of individual limbs or the whole body relative to the ground.

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

2.2 Presence and Task Performance

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

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

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

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

We would nevertheless expect to find an association between presence and performance for some tasks - precisely those tasks that benefit from *immersion*. For the purposes of this study we postulated that increased "immersion" would lead to improved "task performance" (to be defined below in the context of this experiment). This is because the task involved comprehension and memory of a complex three dimensional structure and events relating to that structure, and we considered that performance would be enhanced by an egocentric, stereo view based on a head-tracked HMD compared to an exocentric screen based view. Since our overall hypothesis is that both performance and presence are enhanced by immersion, we would therefore not be surprised to find an association between performance and presence.

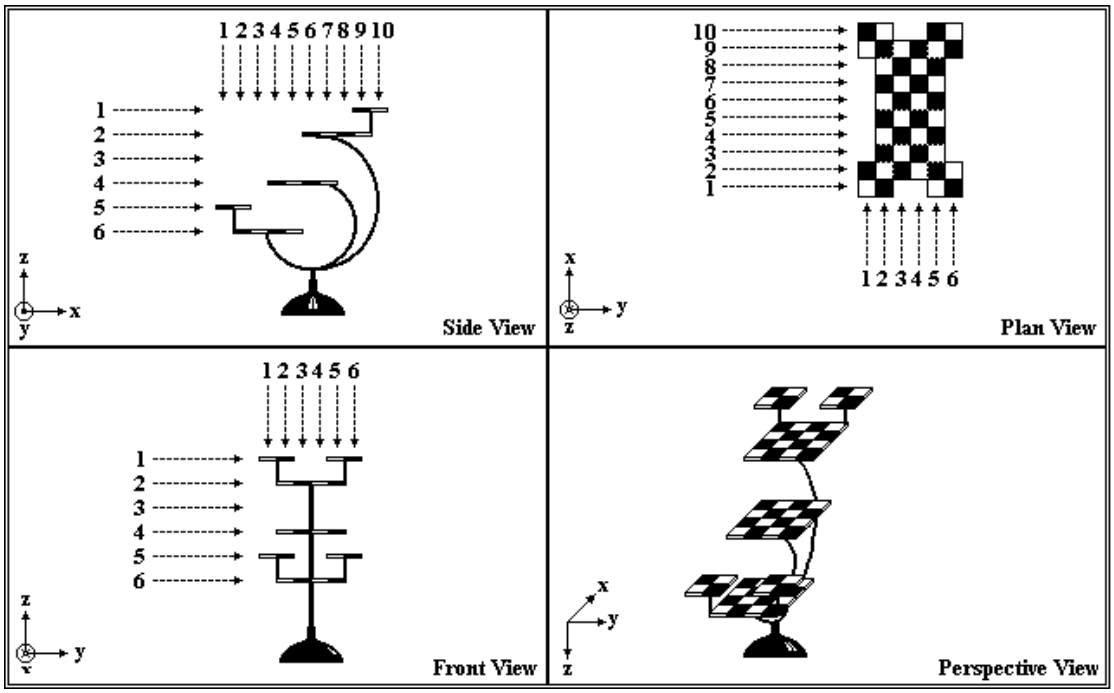

Figure 1
Tri-Dimensional Chess

3. Experiment

3.1 Background: Tri-Dimensional Chess

Tri-Dimensional Chess (TDC) is a board game which has many characteristics in common with conventional chess[3]. More specifically, it is a type of chess played on a number of boards suspended at a different heights. The pieces used in this game are the same as conventional chess and capable of the same movements, but also may be moved from one board to another. Moreover, the layout of the different boards is irregular, and the initial positions of the sixteen pieces of each side are different than in conventional chess. Finally, TDC has a set of four movable attack boards, which are also considered to be pieces, and can be moved according to certain rules (Figure 1).
TDC was chosen because it provides a complex geometrical structure and it is this complexity of the layout of the boards and the pieces, which make it a suitable vehicle for the study. The actual rules of the game were of no importance for this experiment. Twenty four subjects were chosen for the experiment according to the factorial design of Table 1.

3.2 Factorial Design

(a) Immersion: Exocentric/Egocentric

This factor relates to the surrounding aspect of immersion discussed above. Half of the subjects were immersed with an egocentric view into a virtual environment. This was achieved using a DIVISION ProVision100, with a Virtual Research Flight Helmet and a DIVISION 3D Mouse. Polhemus Fastrak sensors were used for position tracking of the head and the mouse. The generated image has a resolution of 704x480 which is relayed to two colour LCDs each with a 360×240 resolution. The HMD provides a horizontal field of view of about 75 degrees, and about 40 degrees vertically. Forward movement in the VE is accomplished by pressing a left thumb button on the 3D mouse, and backward movement with a right thumb button. A virtual hand was slaved to the 3D mouse - there was no virtual body representation other than this. Objects could be touched by the hand and grabbed by using the trigger finger button on the 3D mouse.

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

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

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

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

(b) Environment: Plain/Garden

The environment factor is related to *vividness*. Half of the subjects ("garden") participated in an environment where the TDC system was located in a realistic setting. This consisted of an open field, populated by a table, a chair, a tree and small plant. The TDC board was located on the table. This model had a large horizontal plane forming the ground, and a spherical cone representing the sky. This was called

[3]See http://www.redweb.com/chess/Variants/

the "garden" environment. All surfaces in the VE garden were appropriately texture mapped. The remaining subjects ("plain") saw the TDC game suspended in a void. Examples of these environments are shown in the colour plates.

(c) Number of Moves

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

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

3.3 Virtual Model and Performance

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

The frame rate offered on the ProVision system is not guaranteed at any particular level of performance. It varied between 15 and 20Hz depending on the complexity of the data in view at any particular time. Clearly subjects in the plain environment would have generally experienced a faster frame rate than those in the garden environment. This does confound the experiment to some extent since on the one hand the more realistic environment is, in the terminology of this paper, a more "immersive" one, yet its lower frame rate makes it less immersive. However, an experiment of Barfield and Hendrix (1995) found that frame rates of between 15 and 20Hz resulted in the same degree of reported "presence".

3.4 Procedures

(a) Selection of Subjects

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

(b) The Pre-Questionnaire

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

(c) The Spatial Awareness Test

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

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

(d) Introducing Subjects to the Tri-Dimensional Chess

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

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

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

(e) The Virtual Kitchen Task

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

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

Initially, the VR equipment was shown to the subjects. They were told how they could navigate through the virtual environment, how they could pick objects up, and release them. They were given written instructions on what they had to do in the virtual kitchen. The same instructions were repeated verbally by the experimenter. Each subject had to navigate around the kitchen, find a particular object in the kitchen, pick it up and drop it, after having taken it to another part of the environment. The particular object turned its colour to bright red when touched, hence indicating that it could be picked up. This was done so that consistency would be maintained with the change of colour of pieces in the main experimental task. The subjects were told that objects in Virtual Environments are not solid and are hence penetrable by the virtual hand. They were also made aware of the fact that the virtual environment does not simulate gravity. Finally, the experimenter guided each one of the subjects through this familiarisation process by talking to them and by making himself clearly present in the real world. This was a practice that was avoided during the main experimental task, since talking to subjects during their VR experience is likely to adversely affect the sense of presence. During the main experiment, assistance and guidance was given only in cases of emergency (e.g. when a subject was in danger of colliding with an object in the real world), or in cases where the subject had specifically requested some help.

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

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

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

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

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

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

(g) The Post-Questionnaire

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

- the sense of "being there" in the environment depicted by the display;
- the extent to which there were times that the virtual world seemed more the presenting reality than the real world, and
- the sense of having visited a place rather than seeing images.

3.5 Variables Measured

(a) Response (Dependent) Variables

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

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

(b) Independent Variables

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

(c) Explanatory Variables

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

Table 2
The Main Explanatory Variables

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

4. Results

4.1 Statistical Analysis for Immersion and Presence

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

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

4.2 Statistical Analysis for Performance

Table 3
Mean and Standard Deviations of Proportions of Correct Moves

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

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

Table 4
Logistic Regression of Number of Correct Moves
Overall fit: Deviance = 17.626 on 16 d.f

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

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

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

4.3 Interpretation of Results

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

- Performance as measured in this particular experiment is positively associated with egocentric immersion in comparison with the exocentric screen based viewpoint;
- Performance is positively associated with a more realistic (garden) environment compared with an empty environment;
- Females do not do as well as males in this particular experiment, but
- In the case of females a higher SAT score is associated with increased performance, whereas for males the SAT score is not associated with performance.

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

- Previous knowledge of chess is associated with better performance;
- Better performance is achieved the greater the number of practice sequences.

5. Conclusions

In this paper we have distinguished between immersion and presence, and considered the relationship of these to performance. We have argued that immersion is a description of a technology, whereas presence is concerned with the concomitant behavioural and psychological responses of people. In discussions of "performance" in relation to VEs it should always be clear as to whether the performance relates to efficiency regarding the performance of some task within the VE or in the real world subsequent to a VE experience. We argue that although increased immersion may well improve performance in certain tasks due

to the higher quality and quantity of information available, there is no particular reason to expect presence to improve performance. Presence is concerned with how well a person's behaviour in the VE matches their behaviour in similar circumstances in real life, rather than with how well they perform as such.

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

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

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

Acknowledgements

This work is funded by the UK EPSRC Distributed Extensible Virtual Reality Laboratory project, and Department of Trade and Industry through the London Parallel Applications Centre. Thanks to Cristina Alberto for her AutoCAD help, and to Mark Brown for help with the SAT. We thank Professor Woodrow Barfield of the University of Washington for his helpful comments on an earlier draft.

References

Barfield, W. and Hendrix, C. (1995) The Effect of Update Rate on the Sense of Presence within Virtual Environments, Virtual Reality: The Journal of the Virtual Reality Society, 1(1) 3-16.

Barfield, W. Sheridan, T., Zeltzer, D. and Slater, M. (1995) Presence and Performance Within Virtual Environments, in W. Barfield and T. Furness (eds) Virtual Environments and Advanced Interface Design, Oxford University Press.

Bishop, G. et. al. (1992) Research Directions in Virtual Environments: Report of an NSF Invitational Workshop, Computer Graphics, 26(3), 153-177.

Chung, J. (1992) A Comparison of Head-tracked and Non-head-tracked Steering Modes in the Targeting of Radiotherapy Treatment Beams, Proceedings of the 1992 ACM Symposium on Interactive 3D Graphics, 193-196.

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.

Kendall, M. [1975] *Multivariate Analysis*, Charles Griffin & Co. Ltd.

Linakis, V. (1995) A Study on the Transfer of Skills using a Virtual Tri-Dimensional Chess System, Advanced MSc Thesis, Department of Computer Science, QMW University of London.

Mizell, D., Jones, S., Jackson, P. and Picket, D. (1995) Is VR Better than a Workstation: A Report on Human Performance Experiments in Progress, Eurographics Workshop on Virtual Environments, Proceedings ed. M. Goebel, Monte Carlo, Jan 31-Feb 1st, 1995.

Pausch, R., Shackelford, M.A. and Proffitt, D. (1993) IEEE Symposium on Research Frontiers in Virtual Reality, 41-45.

Slater, M., M. Usoh, A. Steed (1994) Depth of Presence in Immersive Virtual Environments, Presence: Teleoperators and Virtual Environments, MIT Press 3(2), 130-144.

Slater, M., C. Alberto, M. Usoh (1995) In the Building or Through the Window? An Experimental Comparison of Immersive and Non-Immersive Walkthroughs, Virtual Reality Environments in Architecture, Leeds 2-3rd November 1994, Computer Graphics Society.

Slater, M. and Usoh, M. (1995) Modeling in Immersive Virtual Environments: A Case for the Science of VR, in R. Earnshaw, J. Vince, H. Jones (eds) Virtual Reality Applications Academic Press, 53-70.

Slater, M., M. Usoh, A. Steed (1995) Taking Steps: The Influence of a Walking Metaphor on Presence in Virtual Reality, ACM Transactions on Computer Human Interaction (TOCHI) Special Issue on Virtual Reality, September 1995.

Smith, P. and Whetton, C. (1988) General Ability Tests (User's Guide) The National Foundation for Educational Research, ASE.

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

Wilson, P.N., N. Foreman (1993) Transfer of Information from Virtual to Real Space: Implications for People with Physical Disability, in M. Goebel (ed.) First Eurographics Workshop on Virtual Reality, 21-25.

Appendix A

Logistic Regression Model

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

$$\eta_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} \quad (i = 1,2,...,N) \tag{1}$$

where N (=24) is the number of observations. The logistic regression model links the expected value $E(p_i)$ to the linear predictor as:

$$E(p_i) = \frac{n}{1+\exp(-\eta_i)} \tag{2}$$

where n (=3) is the number of binomial trials per observation (the three presence questions). Maximum likelihood estimation is used to obtain estimates of the β coefficients. The deviance (minus twice the log-likelihood ratio of two models) may be used as a goodness of fit significance test, comparing the null model ($\beta_j = 0, j = 1,...k$) 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 χ^2 distribution with degrees of freedom dependent on the number of parameters (added or deleted).

Appendix B

Analysis of Correct Moves

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

$$\begin{aligned}\hat{\eta} &= 0.9535 - 2.477 + 0.5608\times\text{practice} \\ &= -1.5235 + 0.5608\times\text{practice}.\end{aligned}$$

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

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

$$\begin{aligned}\hat{\eta} &= 0.9535+1.127+1.820+0.5608\times\text{practice} \\ &\quad -13.47 + 0.1909\times\text{SAT} \\ &= -9.570 + 0.5608\times\text{practice} + 0.1909\times\text{SAT}\end{aligned}$$