A Direct Comparison of Presence Levels in Text-based and Graphics-based Virtual Environments

David Nunez and Edwin Blake
CVC Laboratory, Department of Computer Science
University of Cape Town
Rondebosch, 7701, South Africa
+27 21 650 2670
{dave,edwin}@cs.uct.ac.za

ABSTRACT

We present results from an experiment (n=78) which used two presence measures to directly compare presence in text-based and graphics-based VEs of two levels of quality. The results show that text-based VEs produce less presence than graphics-based VEs, but the actual difference is less than 20%. This finding has implications for those wishing to implement VEs in impoverished devices or those working towards understanding the cognitive processing of VEs.

CR Categories: I.3.7. [**Computer Graphics**]: Virtual Reality

General Terms: Experimentation, Human Factors, Theory

Keywords: Presence, Priming, Virtual Environments

1. INTRODUCTION

Presence is a phenomenon which can occur when a user interacts with a VR system. It has been described in various ways, ranging from the degree to which users feel they are "there" in a virtual environment rather than in the real environment [1] to the sense that the experience in the virtual environment is no longer mediated by a system [2]. Researchers in the field argue that higher levels of presence can lead to a variety of benefits, such as better task performance or improved navigation in a virtual environment [3]. A great deal of research into the sense of presence in virtual environments already exists, although the great majority of this research focuses on interactive graphical VR systems (see [3], [4], [5], [6] and [9] for examples of this type of work). However, very little presence research has investigated presence in text-based displays such as those used in text-based systems which represent physical spaces (such as MUDs or MOOs). MUDs and MOOs represent locations by means of text descriptions (e.g. "You are on the River Frigid in the vicinity of the Dam"), and navigation and interaction with the system is done by means of text input at a prompt (e.g. "run North" or "light candles with torch"). Some MUDs and MOOs also provide still images of locations in addition to the text.

A body of work in the presence field which suggests that presence can only be expected to occur if the user is presented with high-fidelity displays which allow real-time interaction. For example, Steuer [6] outlines five important variables contributing to presence, namely breadth, depth, speed, range and mapping. Of these, only one (range, which is the degree to which the environment can react to the users' actions) can be implemented to any satisfactory degree in text-based interfaces; the other four all require some degree of sensory fidelity (although it is not yet clear exactly how much). For reasons such as this, text-based interfaces may seem unsuitable for applications where presence is a desired consequence. However, MOOs and MUDs (which it seems at first glance should produce low presence levels) remain extremely popular with users. For instance, the role-playing MUD *Achaea, Dreams for Divine Lands,* has, at the time of writing, an average of 100 players connected on any day, and consists of more that 15,000 rooms [7]. Clearly, although text-based VEs might seem obsolescent, there is still a large group of users who are willing to engage with virtual environments by means of such interfaces. Furthermore, the rise in popularity of hand-held devices with small screens and relatively small processors not powerful enough to display graphics-based VEs (such as mobile telephones and personal data assistants) suggests that being able to produce presence with such small-scale systems is a practical problem which could see widespread application.

We believe that although there is significant evidence (both theoretical and empirical) that presence can be enhanced by increasing the sensory fidelity of graphical systems, there has not, we feel, been sufficient empirical investigation into how presence can be generated by non-graphical or low-fidelity displays. Our interest is not in determining if there exists some relationship between fidelity and presence as a general rule (although this is an interesting question); our interest is in determining if presence occurs for users of non-graphical displays, and if so, we are interested in comparing it to the degree of presence experienced by users of graphical systems.

As presence has been considered by some to be the defining characteristic of VR [8], and because of the theoretical benefits of presence mentioned above, we decided to investigate the sense of presence in text-based displays. Some work has already been done on the sense of presence in virtual environments such as MUDs and MOOs (see [10] and [11] for examples), but there has been very little work on directly investigating differences in the sense of presence between text-based and graphics-based VEs.

Our text system was not intended to represent or approximate the appearance of contemporary MUDs or MOOs, as our interest does not lie in discovering the presence potential of this type of system specifically, but rather the presence generating capacity of text-based systems in a more general sense. The interface we used was in fact based on the interface used by a style of single-player text based adventure games popular in the 1980s, such as The Hobbit [12] or The Pawn [13] (technical details on such systems can be found in [14] which describes the one of seminal games of this genre, Zork). We felt that this type of interface was a suitable choice for our purposes, as this type of game emphasised navigation around a believable space (as does our graphics based system), rather than the more social aims of contemporary MUDs and MOOs. This paper describes our preliminary investigation into the difference in presence levels experienced in text-based VEs as compared to graphics-based VEs.

2. EXPERIMENT

We conducted this experiment as part of a larger study into the cognitive processing of virtual environments. A direct comparison of presence scores collected from three separate groups of users each visiting the same virtual environment (a medieval monastery; cf. 2.2 below for details) was made. Each of the three groups viewed the environment on a different display system: a high quality graphical system, a low quality graphical system, and a text-based system. The purpose of the experiment was to investigate how text displays would compare to graphical displays with regard to presence. We measured presence by means of the Slater, Usoh & Steed presence scale [9], and the Presence Questionnaire of Witmer & Singer [3]. We selected these scales as they are among the most used and most understood measures of presence, and come closest to an accepted standard of presence measurement.

2.1 Participants

We recruited 78 paid student volunteers. Volunteers were mostly first-year science faculty students. Demographic information was not recorded. The participants were randomly assigned into one of three conditions: 27 participants in the *high quality graphics* condition (*H*), 26 into the *low quality graphics* condition (*L*), and 25 into the *text* condition (*T*).

2.2 Virtual Environments

We created our environment as a generic medieval European monastery. The environment was created in three forms; a high quality graphics form, a low quality graphics form, and a text form. The high quality graphics version rendered the scene at 640x480x16 resolution, including textures, radiosity and sound. This system maintained an average frame refresh rate of 17Hz throughout the experiment. The low quality graphics version also rendered the scene at a resolution of 640x480x16, but used flat shaded polygons and no sound. This version also maintained an average refresh rate of 18Hz. The text form (rendered on a 320x240x8 display) used text descriptions of the rooms accompanied by a low resolution (280x100x8) still image. A refresh rate measure is not appropriate for the text based VE, as it was event driven. A comparison of the three versions can be seen in Figures 1, 2 and 3 respectively.

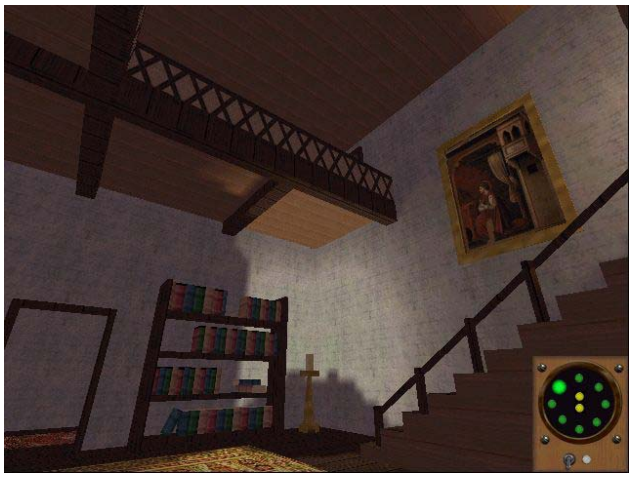

Figure 1: High quality graphics version of the VE

Note that the three versions represent the same building, as they were based on the same floor plans. Both graphical environments could be explored interactively by means of the mouse and keyboard. The text environment was explored by means of key commands, selected from a list of options (such as go North, go South, go up, etc). The list of options adapted itself based on the room the user was in to provide only sensible options (for instance, if there was no exit available to the North, no "go North" option was provided). At any time, the text based system allowed the same actions available in the graphics-based systems.

The VE presented in graphical versions contained 18 furnished rooms divided over three levels, connected by two stairways. The text version presented 27 rooms, 20 of which contained an image accompanying the text. More rooms exist in the text version than in the graphical versions, because stairways and some long passageways are counted as rooms due to the requirements of our text VE system.

Figure 2: Low quality graphics version of the VE

The text version of the VE contained only information which was available in the high quality graphical version. This included

descriptions of the visuals and sound. No information from other sources (e.g. emotions, smells etc) was provided.

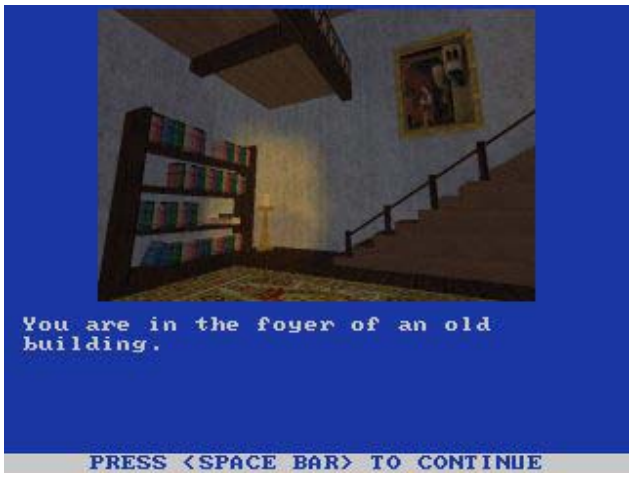

Figure 3: Text version of the VE

2.3 Procedure

The experiment was conducted in three phases. In the first phase, the participants were introduced to the system, instructed in its use and allowed to briefly practice (5 minutes) in a specially prepared training VE. An experimenter was present during the training session to ensure that all participants had learnt to use the system fully before the training session was completed. The second phase had the participants exploring the monastery VE for a timed period of 15 minutes, although the participants were not made aware that a time limit was imposed. To motivate the participants to explore the environments fully, a set of twenty boxes were placed in the virtual environment and the participants were given the task of locating them. This ensured that the space was actively explored, and that the participants' attention was focused on the virtual environment. We did not measure task performance; we simply included the task as a means of keeping the subjects' attention focused on the VE and ensuring that active exploration occurred. In the third and final phase, the participants were led to an adjoining room and were asked to complete questionnaires about their experience. Among these questionnaires were the Presence Questionnaire [3] and the presence scale of Slater, Usoh & Steed [11].

3. RESULTS

Descriptive statistics for the samples are presented in Table 1 and Table 2. The data collected was analyzed using a one-way analysis of variance (ANOVA) to check for differences in the means of each of the conditions. Separate analyses were done for the Presence Questionnaire (PQ) and the presence scale of Slater, Usoh and Steed (SUS). The ANOVAs were statistically significant; the results are presented in Table 3. Further post-hoc tests were done using a series of protected t-tests.

3.1 Post-hoc analysis: SUS

For the SUS, a significant difference was found between the *H* condition and the *T* condition ($t = 2.266$, $df = 50$, $p = 0.02$). A significant difference was also found between the *H* condition and the *L* condition ($t = 2.45$, $df = 51$, $p = 0.017$), but there was no significant difference between the *L* condition and the *T* condition ($t = 0.895$, $df = 49$, $p = 0.375$).

3.2 Post-Hoc analyses: PQ

For the PQ, all differences were significant: the difference between the *H* condition and the *T* condition ($t = 5.2$, $df = 50$, $p < 0.001$), the difference between the *L* condition and the *T* condition ($t = 2.439$, $df = 49$, $p = 0.018$) as well as the difference between the *H* condition and the *L* condition ($t = 2.01$, $df = 51$, $p = 0.048$).

Condition	Mean	Std. dev.	N
High qual gfx	155.33	22.31	27
Low qual gfx	139.92	32.53	26
Text	119.28	27.57	25

Table 1: PQ descriptive statistics

Condition	Mean	Std. dev.	N
High qual gfx	26.66	6.92	27
Low qual gfx	21.07	9.52	26
Text	22.96	4.52	25

Table 2: SUS descriptive statistics

Variable	MS effect	MS error	F(2, 75)	*p*
PQ	8468.531	768.7851	11.01547	.00006
SUS	214.8277	53.424	4.02118	.022

Table 3: ANOVA results

3.3 Discussion of results

The results confirm findings in [10] that text VEs can produce presence to a degree measurable by established presence measures. Of particular interest is the pattern of differences in the two scales. In the case of the SUS, the *T* condition produces presence levels equivalent to the *L* condition. However, this difference may be an artifact of the scale itself, as the difference does not exist in the PQ; the pattern presented by the PQ is more intuitive – the three conditions are significantly different from each other, with the order being *H*, *L* and *T*. This fits into the general hypothesis put forward by Steuer [6] that presence levels vary as a function of the amount of information presented to the user by the display (although it seems that this does not have to be presented as raw sensory information). Most noteworthy in these results is the magnitude of the differences in the means between the various groups. Considering that the SUS consists of 6 items, the difference between the means of the *H* condition and the *T* condition (a difference of 3.7) represents a difference of, on

average, 0.6 which represents less than one single point per question. Similarly, in the PQ (which consists of 32 items), the difference between the means of *H* and *T* represents on average only slightly more than one point per question (1.12 points). As each of these presence scales measures on a scale of 1 to 7, the differences represent an average per-item difference of 16% on the PQ, and 8% on the SUS. Although these differences are statistically significant, their absolute magnitude is quite small, suggesting that although we can expect users of text interfaces to virtual environments to experience less presence than users of graphical systems, the difference in the experience will be small.

4. CONCLUSIONS

The results presented in this paper confirm that graphics-based virtual environments produce statistically higher levels of presence than text-based systems. However, the actual difference produced is quite small, and we propose that this difference can be regarded, for many practical applications, as negligible. The designer of text-based virtual environments can expect the presence on those systems to be lower than in graphics-based systems, but by less than 20%. Although lower presence implies a reduction in benefits such as improved task performance and more effective navigation [3], the fact that the difference is small in turn implies that the reduction of these benefits will also be small.

A criticism which can be leveled at this work is the question of whether the difference in sensory stimulus quality across the three conditions was sufficiently large. Indeed, one might even argue that the small difference in presence levels is due to a small difference in sensory stimulus quality. This is a valid criticism, as we do not provide a measure of stimulus quality; however, we are not aware of any accepted measure for such a construct, and are thus unable to provide such an indication. We feel that this criticism overemphasizes experimental rigour while under-emphasizing ecological considerations. Our choice of system was based on an attempt to understand presence as it occurs currently in field applications. We feel that the graphical systems we used (particularly the high quality graphics system) represents technology which one can expect to find in many applications (such as gaming or desktop visualization). Using a high-resolution head mounted display might have lead to a higher fidelity display and perhaps a more dramatic difference in presence levels, but it would have not provided us with a sense of the differences in presence which one might find in the field.

To conclude, we propose that knowledge of this difference in presence can be useful in calculating tradeoffs between the cost or portability of systems in relation to the level of presence they can produce. This finding is also of interest to those working towards a cognitive theory of presence. The small difference in presence scores between text-based and graphics-based systems suggests that presence is produced not as a result of direct perceptual processing (that is, by “fooling the mind’s eye”, as suggested in [15]), but rather at a higher, more conceptual level. We believe that this evidence strongly suggests that presence could be significantly affected by manipulating cognitive mechanisms other than those involved in perception.

We would like to thank Rudolph Neeser for his assistance in the writing of this paper.

5. REFERENCES

[1] Slater, M. & Steed, A. A virtual presence counter. Presence, 9 (2000), 413-434.

[2] Witmer, B. G., & Singer, M. J. Measuring presence in virtual environments: A presence questionnaire. Presence, (1998), 225-24.

[3] Lombard, M. & Ditton, T. At the heart of it all: the concept of presence. Computer-Mediated Communication, 3, (1997).

[4] Hendrix, C. & Barfield, W. Presence in virtual environments as a function of visual and auditory cues. Proceedings of the Virtual Reality Annual International Symposium ’95, 74-81 (1995).

[5] Usoh, M., Arthur, K., Whitton, M., Bastos, R., Steed, A., Slater, M. & Brooks, F. Walking > walking in place > flying, in virtual environments. Proceedings SIGGRAPH 99, (1999), 359-364.

[6] Steuer, J. Defining virtual reality: dimensions determining telepresence. Journal of Communication, 42, (1992), 73-93.

[7] The Mud Connector. Available Online: www.mudconnector.com

[8] Blake, E., Casanueva, J. & Nunez, D. Presence as a means for understanding user behaviour in virtual environments. South African Computer Journal, 26 (2000), 247-252

[9] Slater, M., Usoh, M. & Steed, A. Depth of presence in virtual environments. Presence: Teleoperators and Virtual Environments, 2, (1995), 201-219

[10] Towell, J & Towell, E. Presence in text-based networked virtual environments or "MUDS". Presence, 6(5), (1997), 590-595.

[11] Schiano, D.J. Lessons from LambdaMOO: a social, text-based virtual environment. Presence, 8(2), (1999), 127-139.

[12] The Hobbit (Sinclair ZX Spectrum version). Melbourne House (1983).

[13] The Pawn (Commodore 64 version). Magnetic Scrolls Inc. (1985).

[14] Lebling, P.D., Blank, M.S., & Anderson, T.A. Zork: A computerized fantasy simulation game. IEEE Computer, 12(4), (1979), 51-59.

[15] Slater, M. Presence and the sixth sense. Presence, 11(4), (2002), 435-43