THE THEMATIC BASELINE TECHNIQUE AS A MEANS OF IMPROVING THE SENSITIVITY OF PRESENCE SELF-REPORT SCALES

David Nunez and Edwin Blake (dave,edwin@cs.uct.ac.za) Collaborative Visual Computing Laboratory Department of Computer Science University of Cape Town, South Africa

Paper presented at the 6th International Conference on Presence, Aalborg, Denmark, October 2003

There has been considerable criticism of the use of self-report scales for the measurement of presence (for instance, see Slater, 1999 and Usoh et al, 2000). During our work with self-report scales (specifically the scales of Slater, Usoh & Steed, 1995 and that of Witmer & Singer, 1998), we have found that the scales are sometimes unable to find differences under conditions which, according to the literature and subjective qualitative reports, differences in presence should exist. For instance, in Johns et al (2000), the authors created two immersion conditions by manipulating nine of seventeen display and interaction factors described by Witmer & Singer (1998) as being determinants of presence. Although the two conditions were qualitatively markedly different, no significant difference was found using Witmer and Singer?s own Presence Questionnaire (1998). Similarly, Usoh et al (1999) created three conditions by altering the mode of locomotion within the VE. But although these three conditions represented an increase in naturalness of interaction and in consistency of VE experiences with the subjects? experiences of moving in the real world, Usoh et al found no difference in presence scores between the three groups (as measured on Slater, Usoh & Steed?s (1995) presence scale). In a more recent example, Mania (2001) used presence questionnaires to measure presence levels of subjects in five display conditions, but found no significant difference between the conditions, even though some of the scores were taken from users experiencing a real environment. A final and perhaps most striking example comes from Usoh et al (2000). They investigated the difference between presence scores derived from subjects in a virtual environment of their laboratory with those of subjects which had visited the actual laboratory; and although the VE was, by the admission of the authors, a low fidelity recreation of the room, no significant difference (using the Slater, Usoh & Steed (1995) presence scale) was found between this group and those who had visited the real laboratory.

We believe that a major reason why self-report scales often fail to detect differences in situations where presence differences probably exist is that many items found in current presence questionnaires require the subjects to compare their experiences in the virtual environment with other experiences they are familiar with. Consider, as an example, the following item from Slater, Usoh & Steed?s (1995) scale (we use this as an example mostly because it is a well known and widely used scale):

To what extent were there times during the experience when the virtual environment was the reality for you?

Another item from the same scale:

Please rate your sense of being in the virtual environment, on a scale of 1 to 7, where 7 represents your normal experience of being in a place.

These items require the subject to make a comparison between their VE experience and some standard ?real? experience, and then to assign a value indicating degree to which the experiences matched. However, the standard experience against which the subjects are to compare is not defined, and each subject is thus left to interpret the task posed by the item in their own way. This leads to an increase in the randomness of response, and thus an increase in the error variance of the scale, which in practice means a decrease in the sensitivity of the scale (Anastasi & Urbina, 1996).

It seems to us that the sensitivity of self-report scales could be improved by providing subjects with a standard or semi-standard baseline against which they could gauge their VE experiences. Finding an appropriate baseline is however quite problematic. An explicit baseline based on image quality, sound spatialization or other display variables is problematic due to its blurring of the boundaries between immersion and presence (see Slater, 1999 and Singer & Witmer, 1999 for more on this problem). Ideally, the baseline should allow subjects to compare in terms of their own experiences, while still allowing the expression of individual differences in presence scores.

Because the VEs to which users are exposed are often beyond their own experience (particularly in entertainment applications such as games or in training applications such as flight simulators where the user has not yet experienced the real setting), we propose using the user?s natural expectations as a measurement baseline. To operationalize the concept of natural expectations, we make use of the schemata and script concepts used in

cognitive psychology (see Schank & Abelson, 1977 and Rumelhart & Ortony, 1977), which encode mental representations of events or stories. Thus, in a VE of Elizabethan England, the expectations of what should occur and what is ?natural? or ?realistic? in such a setting would be encoded by a set of schemata and scripts related to Shakespeare, powdered wigs, renaissance times, and so on. It is worth noting that although the subject?s culture plays a role in shaping schemata and scripts, giving them some uniformity, there will be individual differences, so the expectations in a setting will vary slightly from person to person, much in the same way that the experience of visiting a real place will vary from person to person. Using the concepts of schemata and scripts has the added advantage that previous research has shown that they can be primed or activated by presenting a subject with various materials such as texts or images (Rumelhart & Ortony, 1977). Thus, by showing a subject pictures or a texts to do with Elizabethan England, the expectations of that subject can be manipulated so that the cognitive processing of material related in theme would occur morereadily than the processing of material with a different theme (Schank & Abelson, 1997).

We believe that by priming a subject with materials related in theme to the VE they are about to experience, relevant schemata and scripts will become activated, effectively providing a subjective and yet consistent baseline against which subjects can gauge their VE experiences. We conducted an experiment in which we tested the value of this thematic priming in presence measurement, and empirically tested the notion that the thematic baseline provided by priming can increase the sensitivity of measurement

EXPERIMENT

We tested the thematic baseline technique by creating two priming conditions. In the first, subjects were shown priming materials related in theme to the VE before immersion in the VE, and in the other, the subjects would be shown a control set of priming materials (not thematically related to the VE) before immersion. In each of these conditions, two VEs were used, one more rich and complex in terms of visual and auditory stimuli than the other. This was done to create an average difference in presence scores between the VEs. We based our expectation of a difference in presence scores under these conditions based on the suggestions by Steuer (1992), Zeltzer (1992), Bystrom et al (1995), Sheridan (1996) and others, who suggest that presence is related to the quality of the display, and to multimodality in displays. Our design was a 2x2 factorial ANOVA design. The factors were stimulus quality of VE display x thematic priming. Our experiment used a desktop-based system displaying a 640x480x16 graphical stream at an average of 15Hz, and stereo sound played on headphones. The stimulus quality condition was manipulated by creating two forms of VE: a high quality version which included textures, radiosity and 3D sound, and a low quality version which used flat shaded polygons and no sound.

To manipulate the thematic priming variable, we activated the relevant schemata in our subjects by using printed booklets, containing descriptive texts and images. Each subject read one of two booklets before immersion in the VE: either one related thematically to the virtual environment (the VE relevant priming condition) or one not thematically related to the VE (the VE irrelevant priming condition). The booklets were read immediately prior to immersion in the VE.

We measured presence by the use of the the Slater, Usoh & Steed (1995) questionnaire (SUS) and the Witmer & Singer (1998) Presence Questionnaire (PQ). Two scales were used so that the generality of the technique in relation to presence scales could be established.

RESULTS

We collected a total of 103 sets of observations (a set constitutes an SUS and PQ score). The observations were assigned to one of the four cells of the design randomly. The data were analyzed using the factorial analysis of variance (ANOVA) technique.

The SUS showed a significant interaction effect between stimulus quality and priming ($F(1, 99) = 10.18$ $p < 0.0019$). There was also a significant main effect on stimulus quality ($F(1,99) = 9.64$ $p < 0.002$). There was, however, no significant main effect on priming ($F(1,99) = 0.17$, $p > 0.65$). The PQ showed a similar pattern of results to the SUS. The interaction between stimulus quality and priming was significant ($F(1,99) = 4.23$ $p < 0.05$), as was the main effect of stimulus quality ($F(1,99) = 5.99$ $p < 0.02$). The main effect of priming on PQ was not significant ($F(1,99) = 0.23$ $p > 0.63$). A post-hoc investigation of the interaction effects shows a similar pattern for both scales. In both cases, there exists a significant difference in the mean presence levels between the high and low stimulus quality conditions when the subjects were primed with material thematically related to the VE (PQ: $t = 3.42$, $df = 49$, $p < 0.0013$; SUS: $t = 3.99$, $df = 49$, $p < 0.0002$), but no significant difference between the stimulus quality conditions when the subjects were not primed with thematically relevant materials (PQ: $t = 0.26$, $df = 50$, $p > 0.79$; SUS: $t = 0.069$, $df = 50$, $p > 0.945$).

DISCUSSION

The results show that for both measures, a significant difference in presence scores exists between the high and low quality stimulus quality conditions, but only when the subjects were primed with thematically relevant material. This suggests that providing an appropriate cognitive context by means of priming can increase the sensitivity of self-report scales. Our use of two separate measures, and the replication of the finding in both scales (particularly scales which are as different as the SUS and PQ) allows us to suggest that the finding is not due to any idiosyncrasies of a particular scale, but rather a general effect.

This finding has, we believe, particular significance for presence theory, as it implies that presence is not only related to low level perceptual processes, but also partly determined by higher-level conceptual states (as suggested by Slater, 2003). Although our study focused on the effect that priming has on scale self-report scale response, it is more than likely that priming plays a role in the presence experience itself, and that VE engineers may be able to use conceptual and thematic manipulations of subjects to create better presence experiences.

REFERENCES

- Anastasi, A., & Urbina, S. (1996). Psychological Testing. New York, NY: Prentice Hall.
- Barfield, W., Zeltzer, D., Sheridan, T.B., & Slater, M. (1995). Presence and performance within virtual environments. In W. Barfield & T. Furness (Eds.), Virtual Environments and Advanced Interface Design. London: Oxford University Press.
- IJsselsteijn, W. A., de Ridder, H., Freeman, J., & Avons, S. E. (2000). Presence: Concept, determinants and measurement. Proceedings of the SPIE, Human Vision and Electronic Imaging V, pp. 3959-76.
- Johns, C.L., Nunez, D., Daya, M., Sellars, D., Casaneuva, J.S., & Blake, E.H. (2000), The interaction between individuals' immersive tendencies and the sensation of presence in a virtual environment. In J.D. Mulder & R. van Liere (Eds). Virtual Environments 2000: Proceedings of the Eurographics Workshop. Amsterdam, The Netherlands.
- Mania, K. (2001). Connections between Lighting Impressions and Presence in Real and Virtual Environments. Proceedings of the 1st International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa, pp. 119-123
- Rumelhart, D.E., & Ortony, A. (1977). The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro & W.E. Montage (Eds.), Schooling and the acquisition of knowledge. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Schank, R. & Abelson, R. (1977). Scripts, plans, goals and understanding. Hillsdale, NJ.: Erlbaum
- Sheridan, T.B. (1996). Further musings on the psychophysics of presence. Presence: Teleoperators and Virtual Environments, 5, pp. 241-246
- Singer, M.J. & Witmer, R.G. (1999). On selecting the right yardstick. Presence: Teleoperators and Virtual Environments, 8, pp. 566-573
- Slater, M. (1999). Measuring presence: a response to the Witmer and Singer presence questionnaire. Presence: Teleoperators and Virtual Environments, 8 , pp. 560-565.
- Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. Journal of Communication, 42, 73-93.
- Usoh, M., Arthur, K., Whitton, M.C., Bastos, R., Steed, A., Slater, M. & Brooks, F.P. (1999). Walking > walking-in-place > flying, in virtual environments. Proceedings, SIGGRAPH99, pp. 359-364
- Usoh, M., Catena, E., Arman, S. & Slater. M. (2000) Presence questionnaires in reality. Presence: Teleoperators and Virtual Environments, 9, pp. 497-503.
- Witmer, B.G., & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. Presence: Teleoperators and Virtual Environments, 7, pp. 225-240
- Zeltzer, D. (1992). Autonomy, interaction and presence. Presence: Teleoperators and Virtual Environments, 1, pp. 127-132.