Evaluating Relative Impact of Virtual Reality System Variables on Architectural Design Comprehension and Presence

A variable-centered approach using fractional factorial experiment

Loukas N. Kalisperis[1], Katsuhiko Muramoto[1], Bimal Balakrishnan[1,2], Dragana Nikolic[1], Nevena Zikic[1]
[1]Immersive Environments Lab Research Group, Dept. of Architecture, The Pennsylvania State University, United States of America
[2]Media Effects Research Lab, College of Communications, The Pennsylvania State University, United States of America
[1] http://www.arch.psu.edu and [2] http://www.psu.edu/depts/medialab

The relative contributions of five variables (Stereoscopy, screen size, field of view, level of realism and level of detail) of virtual reality systems on spatial comprehension and presence are evaluated here. Using a variable-centered approach instead of an object-centric view as its theoretical basis, the contributions of these five variables and their two-way interactions are estimated through a 2^{5-1} fractional factorial experiment (screening design) of resolution V with 84 subjects. The experiment design, procedure, measures used, creation of scales and indices, results of statistical analysis, their meaning and agenda for future research are elaborated.

Keywords: *Virtual reality system variables - stereoscopy; screen size; field of view; level of realism; level of detail; spatial comprehension; presence; variable-centered approach; fractional factorial experiment design.*

A variable-centered approach to studying the effects of virtual reality technology for architectural design education

Choosing an appropriate medium of representation is critical for developing the ability to visualize space in beginning students of architectural design. Our recent work encourages undergraduate architecture design students to engage in dynamic 3-D spatial analyses throughout the design process in the ITS/SALA Immersive Environments Laboratory (IEL). The IEL offers a large three-screen, panoramic, stereoscopic virtual reality (VR) display allowing students to visualize and evaluate their architectural design projects throughout the design process. Our experiences at the IEL (Kalisperis et al., 2002) and those of others have validated the importance of VR for architectural visualization. When we take a holistic view of virtual reality technology comprising of a number of component variables, each having a particular value, it is difficult to determine the relative contributions of specific VR system attributes such as screen size or stereoscopy towards spatial comprehension. An alternative would be to take a 'variable-centered' approach put forth by Nass and Mason (1990) which breaks down the technology into its component variables and their corresponding values. Methodologically, fractional factorial experiments are probably best suited for empirical assessment using a variable-centered approach.

Technological and content specific variables of virtual reality technology

Our earlier studies have validated the desirability of key VR attributes, such as screen size, wide field of view and stereoscopic projection for architectural design presentation (Kalisperis et al., 2002). We have also noticed that the visualizations in the IEL grow more complex and become more realistic as the designs evolve through the course of the studio. Our research elaborated below aims to identify the relative contribution of the following virtual reality system variables on spatial comprehension, drawing on our observations in the IEL and review of literature.

Display variables - Stereoscopy, screen size and field of view

Stereoscopy is perhaps the most important characteristic of any virtual reality system and the case for its inclusion seems commonsensical. Stereoscopy enhances perception of three-dimensional objects on a computer screen through binocular disparity (Hubona et al., 1997) and thus is critical for spatial visualization. Hendrix and Barfield (1996) point out that stereoscopy adds a sense of presence in addition to the perception of depth and volume. Though there is some debate in the literature as reviewed by Hubona et al. (1997) about the positive effect of stereoscopy on task performance in different domains, it seems reasonable to assume that overall it can have a positive impact on perception of space.

Large displays provide a greater sense of presence and Tan et al. (2003) have found that the larger displays help users treat spatial tasks as egocentric as opposed to exocentric, resulting in better presence and improved performance on spatial tasks. Patrick et al (2000) offers better allocation of attentional resources as another explanation for improved task performance using large displays in addition to the egocentric locus and presence suggested by Tan et al (2000).

Another variable related to screen size that can potentially influence spatial comprehension and presence is the field of view. Prothero et al. (1995) found that wider field of view enhances ones' feeling of presence. Comparing desktop monitor, large display and head mounted display, Patrick et al. (2000) found that large projection displays covering the entire field of view helped viewers form better cognitive map. Citing prior research, Arthur (2000) points out that narrow field of view in a real world lowers human performance on navigation, spatial awareness, coordination and perception of size and space. Similarly in a virtual environment, a narrow field of view makes object appear nearer (Arthur, 2000). Czerwinski et al. (2002) suggest that narrow field of view deteriorates perceptual, visual, and motor performance in various tasks and therefore wider field of view afforded by large displays can improve performance in three dimensional navigation tasks.

Content variables - Level of Detail and Level of Realism

Visualizations during design process vary from highly abstract representations in the early stages to highly detailed and realistic ones in the closing stages. We are interested in understanding how these content factors can impact the spatial perception and how it might interact with screen size, field of view and stereoscopy. While *level of realism* and *level of detail* commonly to refer to the same idea, we define the two distinctly. The goal of most virtual reality systems is representational and functional isomorphism with a corresponding real world scenario (Otto, 2002). While functional isomorphism is usually achieved through navigational and behavioral constraints, representational isomorphism is achieved through photorealistic rendering of the virtual world, complete with textures, real world lighting, shade and shadows and color all of which act as cues in perceiving spatial depth according to Michel (1996). The term realism in this study refers to photorealism. Witmer and Singer (1998) suggest that greater pictorial realism increases the sense of presence. Low and high levels of realism were decided through pre-testing. In high realism we included textures, lights, shadow, shade, reflection, refraction, environmental maps. In low realism, we replaced textures with plain colors, and removed reflection, refraction, shadows, and environmental maps.

As design evolves, designers progressively populate their visualizations with more functional elements or details such as furniture, which not only add to realism but act as depth cues. Understanding the role of these cues is important, since mistakes in this regard can distort the spatial perception (Michel, 1996). We classified these elements into four categories. Basic furniture elements such as dining table or bed, which help understand and evaluate the function of the space, formed the most elementary category. Standardized fixtures such as doorknobs, light switches etc that could help determine the size and scale of spaces by virtue of their fixed sizes and location with respect to the human height constitute the next important category. Third category included furniture elements that can further contribute to ones spatial perception though not as crucial as the ones in the first category. The last category comprised of purely decorative elements such as plants, which might contribute to a greater sense of presence. Through pre-testing it was decided to keep all four categories in the high condition and in low condition we discarded fixtures and decorative elements.

Method

Fractional factorial (screening) experiment design

A fractional factorial experiment design is appropriate for a variable-centered approach to studying technology, since its cost-benefit ratio is more efficient. Our goal was to identify any main effect and lower order interaction effects that can affect spatial comprehension and presence, while keeping the number of experimental units, the number of subjects and time required within reasonable limits. The experiment was therefore planned as a 2^{5-1} design requiring 16 experimental units. The highest order interaction, i.e. the 5-factor interaction *screen size*field of view*level of detail*level of realism*stereoscopy* was used as the defining contrast to generate the design so that none of the main effects or 2-factor interactions is confounded with any other 2-factor interactions or main effects, resulting in a design of resolution V.

Participants & Procedure

Participants (N=84) were drawn from second through fifth year design studios of the undergraduate program in the department of Architecture at Penn State University. The average age of the subject was 21.5 years (S.D.=1.75) and there were equal number of male and female participants.

Participants were randomly assigned to one of the sixteen experimental conditions, ensured that they had normal or corrected to normal vision and those in the stereo condition were tested for stereo blindness. Stimulus based on the experimental condition was presented to the subject and the responses to the presence part of the questionnaire were collected. The same stimulus was again shown in five segments corresponding to exterior, living room, dining room, kitchen and study room. The stimulus was paused at the end of each segment and the subjects completed the distance estimation tasks corresponding to the space they just saw. The subjects then sketched the layout out of the rooms on a given outline, mapped the path of their movement in the first floor and completed questions pertaining to the location of various rooms with respect to other rooms. The last part of the questionnaire dealt with computer use, prior experience with VR facility and demographic information.

Stimulus material

A six minute long walkthrough of a two-storey residence was used as stimulus for the experiment. Depending on the experimental condition, the level of detail and level of realism were either high or low based on the operationalization described earlier and the stimulus presented in stereo or non-stereo. Screen size was manipulated by presenting the stimulus on either 19″ desktop monitors or on large 8′ x 6′ rear projection screens. For the wider field of view, a 3:1 display ratio was used (presented on 3 screens) as opposed to 4:3 (single screen) for the narrow condition. Although navigability is a key feature of virtual environments, a predetermined walkthrough was used to improve precision by ruling out individual differences. Figure 1 through figure 4 shows stimuli in combination of different levels of realism and detail

Dependent measures, scale and index construction

Presence

Presence is synonymous with the virtual reality experience and speculated to improve spatial comprehension. We measured presence using a 13-item, 8-point Likert type scale adapted from the *Igroup Presence Questionnaire* (IPQ) by Schubert et al. (2001), the Witmer and Singer's (1998) *Presence Questionnaire*

Figures 1 to 4
Combination of varying levels of realism and detail for the living room used in the stimuli

(PQ) and the *Reality Judgment and Presence Questionnaire* by Banos et al. (2000). The items measured ones' sense of immersion, the mental effort needed and congruence of the walkthrough experience with reality.

Principal component analysis was used to analyze the dimensionality of the thirteen items used to measure presence. Using the criteria of eigenvalues greater than or equal to one, three underlying factors were identified accounting for 63.23% of the variance. On rotation using a Varimax procedure, seven items loaded clearly onto the three factors with their highest loading exceeding 0.6 and the other loadings less than 0.4. The remaining six items, which cross-loaded across the factors, were discarded from further analysis. The rotated solution yielded three factors, level of immersion, ease of immersion and experiential congruence with real world. The last factor consisted of only one clear loading item and was dropped from further analysis. Items for level of immersion measured the extent to which the subjects felt "they were in the house", the extent to which they felt they could "grasp an object in the house" and how "real" those objects felt. Items for ease of immersion measured the "ease of getting used to the house", the "ease of getting a good feel of the spaces" and the "extent of mental effort required for the experience". These two indices created by averaging the three respective items is more meaningful for further analysis than using a presence index created by additively combining the original 13 items, which still would have good reliability (Cronbach's alpha = 0.82).

Spatial comprehension

Spatial comprehension can be defined in elementary terms as ones' understanding of a given space. We operationalized spatial comprehension as ones' understanding of the dimensions and proportion of the various spaces, their scale or in other terms their relationship to the human body, their location and orientation and their relationship with other spaces in terms of their relative sizes and position. After being taken through each space, the subjects were asked to estimate its width and depth, measure heights of objects or spaces in the scene and distance between different objects. For all open-ended distance estimation tasks, overall depth, width and height scores were computed. Each response was standardized by dividing it with the correct response, since the distances that were to be estimated varied considerably. Thus a value of 1.0 indicates an accurate estimation, value above 1.0 indicates over-estimation and less than 1.0 indicates under-estimation. All standardized responses for the depth questions were averaged to create an *overall depth score*. Similar scores were created for *overall width and overall height*, which were used for final analysis.

To determine the subjects' perception of space with respect to their own body, the tasks included estimating the number of steps required to walk from one point to another, guessing how many people a space can accommodate etc. Knowledge of spatial organization and ability to retrace ones' movement through the space is an important element of spatial comprehension. The subjects were asked to sketch the layout of various rooms on the first floor given the exterior out line and further map their movement path. We coded these for accuracy of room locations, size proportions and the precision of the movement path. Two independent coders scored the responses based on pre-determined coding criteria and they had an agreement of 86%.

Control measures and Manipulation check

The impact of the independent variables was assessed more accurately by partitioning out the amount of variability accounted for by extraneous variables. Some variables such as distance from the screen were maintained at constant levels. Demographic factors such as age, gender, height, academic major, academic standing and previous experience with the experiment facility were measured for statistical control.

To ensure that our operationalization of low and high levels of detail was effective, we included a manipulation check in the questionnaire where the subjects rated how photorealistic and how well furnished various spaces were. An independent sample t-test assuming unequal variances was significant, $t(81) = -5.53$, $p<.01$ confirmed that subjects in the high detail condition perceived the stimulus as more detailed (M= 5.43) compared to the low detail condition (M=4.09). Similar significant test, $t(80) = -6.19$; $p<.01$ confirmed that the participants in the high realism condition perceived the stimulus as more photorealistic (M=5.43) compared to the low realism (M=4.00).

In summary, the final analyses were performed with the following dependent variables - level of immersion, ease of immersion, overall depth score, overall width score, overall height score, spatial organization score, proportioning score and way finding score.

Results and discussion

Analyses of Variances (ANOVA) were conducted to answer the primary research question of how the virtual reality system variables affect presence and spatial comprehension. Wherever it was conceivable that one of the control measures could make a difference, they were added as covariates and Analysis of Covariance (ANCOVA) was performed instead of ANOVA to improve the accuracy of the findings. Since the covariates are not the focus of the analysis, the results for the covariates are not reported here. We have reported results with p-value less than 0.1 instead of 0.05, which is usually the norm since we are more interested in trends, this being a screening experiment.

For *level of immersion*, controlling for the extent of computer use, the analysis of covariance found a significant interactive effect between the field of view and detail, $F(1,65) = 5.41$, $p<.05$, between screen size and stereoscopy, $F(1,65)=4.11$, $p=0.1$ and also between detail and stereoscopy $F(1,65)=6.05$, $p<.05$. For low level of detail, smaller field of view has a greater level of immersion (M=5.37) compared to wider field of view (M=4.90) whereas for higher level of detail, the level of immersion is greater for wider field of view (M=5.21) with respect to smaller field of view (M=4.40). For smaller screen size, there is hardly any difference in level of immersion between stereo (M=5.44) and non-stereo (M=5.29). However, for larger screen size, the level of immersion is much greater for stereo condition (M=5.56) when compared to non-stereo condition (M=4.47). For low level of detail, there is little difference in level of immersion between non-stereo (M=5.29) and stereo (M=5.44) condition whereas for the higher level of detail, the level of immersion is greater for non-stereo condition (M=4.91) compared to stereo condition (M=3.90).

For *ease of immersion*, controlling for the extent of computer use, there was a significant interactive effect between screen size and field of view, $F(1,63)=3.27$, $p<0.1$and also between field of view and realism, $F(1,63)=3.37$, $p<0.1$. For larger screen size, there was hardly any difference in the ease of immersion between smaller (M=5.80) and wider field of view (M=6.05), however for smaller screen size, the ease of immersion was greater for the wider field of view (M=6.61) compared to the smaller field of view (M=5.55). For wider field of view, the ease of immersion was almost the same for low (M=6.61) and high realism (M=6.74) conditions whereas for the narrower field of view, the ease of immersion was much greater for the high realism condition (M=6.50) compared to the low realism condition (M=5.55)

Controlling for subjects' academic standing (year of study), analysis for *overall depth score* revealed significant interactions between screen size and level of detail, $F(1,64)=4.49$, $p<0.05$ and between field of view and level of realism, $F(1,64)=7.80$, $p<0.01$. For the high level of detail, there is hardly any difference in depth perception between smaller (M=1.20) and larger screen size (M=1.22) whereas for the low level of detail, those in larger screen condition (M=1.16) tended to estimate the depth more accurately than those in the small screen condition (M=1.25). For smaller field of view, there is hardly any difference in depth perception between those in the low level of realism (M=1.25) and high level of realism (M=1.24). However for wider field of view, those in high level of realism (M=1.37) tend to greatly overestimate the depth compared to those in the low level of realism condition (M=1.20).

For *overall width score*, controlling for the subjects' academic standing, there was a significant interactive effect for field of view and realism, $F(1,63)=9.28$, $p< 0.05$. For low level of realism, smaller field of view results in greater overestimation of overall width (M=1.24) compared to wider field of view (M=1.16). Where as, for high level of realism, the opposite holds true with wider field of view resulting in greater overestimation (M=1.28) compared to smaller field of view (M=1.16).

For *overall height score*, controlling for the subjects' academic standing, there was a significant interaction between screen size and stereoscopy, $F(1,60)=5.49$, $p<.05$ as well as between level of detail and stereoscopy, $F(1,60)=3.73$, $p<0.1$. For smaller screen size, the overall height estimation was more or less accurate for both stereo (M=1.11) and non-stereo condition (M=1.10) whereas for the larger screen size, those in non-stereo condition (M=1.23) tend to overestimate the overall height compared to stereo condition (M=1.13). For non-stereo, there was no difference between low and high detail (M=1.10) in estimating overall height. However, for the stereo condition, high level of detail tends to result in overestimation of overall height (M=1.20) compared to low detail (M=1.11).

Controlling for the subjects' experience with computer graphics and academic standing, there were significant interactions between screen size and stereoscopy, $F(1,60)=4.04$, $p<0.05$ and also between level of detail and stereoscopy, $F(1,60)=3.43$, $p<0.1$ for spatial organization score. For smaller screen size, there was little difference in understanding spatial organization between stereo (M=8.22) and non-stereo condition (M=7.95) whereas for the larger screen the score was much lower for the non-stereo condition (M=6.46) compared to stereo condition (M=8.44). For low level of detail, there was little difference in spatial organization score between stereo (M=7.95) and non-stereo condition (M=8.22). For high detail condition, the score was much lower for the stereo condition (M=6.30) compared to non-stereo (M=8.62).

There were no significant findings for the proportioning score or way finding score in the analysis suggesting that there was little difference between the experimental conditions in their influence on the two scores.

Limitations and directions for future studies

With such an elaborate dataset, many more relationships were explored, but not included due to constraints of space. This study may have failed to detect some significant effects due to the lack of sufficient statistical power. The second wave of data collection would mitigate this drawback. With enough statistical power, we also plan to investigate the role of presence more closely and if indeed it acts as a mediator between the system variables and spatial comprehension.

Follow up studies intend to build on the findings of this one by eliminating non-significant factors and adding other potentially interesting factors such as screen resolution, luminosity and object-centric or viewer-centric representation. Full factorial experiments with more levels for each factor would be conducted once we identify important factors. Another direction of exploration aims to identify the differences between students of design majors and non-design majors, among different design majors and between student designers and professionals so that VR can be used to bridge any gap that may exist between them.

Acknowledgements

The authors gratefully acknowledge Dr. John Messner for use of the Immersive Construction Lab (iCon

Lab), George Otto & Gavin Burris of the Visualization Group, ITS for the technical support, Dr. S. Shyam Sundar from the College of Communications for guidance on experiment design and the undergraduate students for the participation in the experiment. For further information regarding this study, please contact Dr. Loukas Kalisperis at lnk@psu.edu.

References

Arthur, K.W.: 2000, Effects of field of view on performance with head-mounted displays, Doctoral Dissertation, Department. of Computer Science, University of North Carolina, Chapel Hill.

Banos, R.M., Botella, C., Garcia-Palacios, A., Villa, H., Perpina, C. and Alcaniz, M.: 2000, Presence and reality judgment in virtual environments: A Unitary construct?, Cyberphychology & Behavior, Volume 3, Number 3.

Czerwinski, M., Tan, D.S., and Robertson, G.G.: 2002, Women take a wider view, Proceedings of CHI 2002, pp. 195-202.

Hendrix, C., and Barfield, W.: 1996, Presence within virtual environments as a function of visual display parameters, Presence, 4(3), pp. 274-289.

Hubona, G., Shirah, G. and Fout D.: 1997, The Effects of motion and stereopsis on three-dimensional visualization, International Journal Human-Computer Studies. 47, pp. 609-627.

Kalisperis, L., Otto, G. Muramoto, K. Gundrum, J., Masters, R., & Orland, B.: 2002, Virtual reality/space visualization in design education: the VR-desktop initiative, Proceedings of eCAADe2002 - Design e-ducation: Connecting the real and the virtual, Warsaw, Poland, pp. 64-71.

Michel, L.: 1996, Light: The Shape of Space, New York : Van Nostrand Reinhold.

Nass, C. and Mason, L.: 1990, On the study of technology and task: a variable-based approach, in Fulk, J. (Ed.), Organizations and communication technology, Newbury Park, CA: Charles Steinfield, pp. 46-67.

Otto, G. H.: 2002, An integrative Media Approach to Virtual Reality System Design and Assessment, M.A. Telecom Paper, The Pennsylvania State University, USA

Patrick, E., Cosgrove, D., Slavkovic, A., Rode, J.A., Verratti, T., and Chiselko, G.: 2000, Using a large projection screen as an alternative to head-mounted displays for virtual environments, Proceedings of CHI 2000, pp. 478-485.

Prothero, J.D. and Hoffman, H.D.: 1995, Widening the field of view increases the sense of presence within immersive virtual environments, Technical Report, University of Washington, Seattle, WA, R-95-4.

Schubert, T., Friedmann, F. and Regenbrecht, H.: 2001, The Experience of presence: factor analytic insights, Presence: Teleoperators and Virtual Environments, 10(3), pp. 266-281.

Tan D., Gergle D., Scupelli P. and Pausch R.: 2003, With similar visual angles, larger displays improve spatial performance, Carnegie Mellon University, Pittsburgh, PA, CHI 2003 New Horizons, Ft. Lauderdale, Florida, USA, pp. 217-224.

Witmer, B. and Singer, M.: 1998, Measuring presence in virtual environment: A presence questionnaire, Presence: Teleoperators and Virtual Environments 7(3), pp. 225-240.