VEHanoi: Performance differences in Virtual vs Real Environments

Tyler Waite
School of Library and Information Science
Indiana University

ABSTRACT

This research used a real and a virtual Tower of Hanoi puzzle to examine performance differences between virtual and real environments. The puzzle was solved twice in the real environment and twice in the virtual environment by each of the participants. A one-way ANOVA was performed for each of the four trials to compare the performance difference between the two environments. Comparisons were made on number of moves to solve the puzzle, time to move an object, and wand selection method. The analysis showed a significant difference in favor of real world performance of the task with respect to time to move an object. The paper concludes with a report on participant comments and an exploration of the possible sources of this difference.

Keywords
CAVE, virtual environment, Tower of Hanoi, spatial cognition

INTRODUCTION

Quoting the words of researcher J. Findlay, speaking about virtual environments in general, "These environments may provide the necessary tools to assist researchers in disclosing the mysteries of human perception." (Stanney 1998). Virtual environments (VE) present the researcher with a three dimensional world that is programmable and under their control. Within this computerized world, a person can free their mind of concern about incurring serious physical pain as a consequence of their actions. They can step off the side of a virtual cliff without the fear of death. The finality of actions, which once performed can never be undone, does not apply in this world of light and illusion.

Commenting on the differences between the typical computer screen interface and the experience of a VE, Johnson (1994) states "Virtual reality gives the user an environment to work in, rather than a screen to look at." Researchers in chemistry (Burdea, Richard & Coiffet 1996), meteorology, astronomy, network administration (Cruz-Neira et al 1993, see also Larijani 1994), and air traffic control (Heim, 1993), not to mention psychology (Cruz-Neira et al 1993), just to name a few, have all been exploiting the potential application of virtual reality technology to develop new and exciting environments to work in.

The fact that VEs are creatable provides scientists with a powerful new tool with which to explore human behavior and cognition. With VE's the cognitive scientist can develop a variety of highly controlled scenarios that can be called up in an instance. The researcher can quickly switch the participant in or out of different test scenarios with a few key-strokes. The whole environment can be altered or only one characteristic of one specific object

VE technology encompasses a variety of equipment. There are head mounted displays, monocular displays, augmented reality displays, haptic (tactile) interfaces, desktop 3D worlds, and CAVE environments to name a few. This study used a CAVE environment. The CAVE at Indiana University is comprised of four projection surfaces; three 8 foot by 8 foot screens, and the floor (see fig. 1). The three wall screens use rear projection, while the floor image is produced from an overhead mirror projecting the image on the floor. The screens are arranged in front of and to either side of the person inside the CAVE. The stereoscopic images of the CAVE are produced using shutter glasses which blend the images on the projection surfaces into a virtual world whose contents when approached appears to exist not on the screens but in the open space between the projection surfaces. The glasses also have a head-tracking device connected to them so that the proper perspective and environment view can be presented

For this experiment the participants interacted with the environment using a CAVE wand. The wand is roughly the size of a typical television remote control. The base of the wand is cylindrical to permit easy grasping and holding. The top of the wand has three buttons and a mini joystick for moving around in the virtual environments.

What advantages does a CAVE environment offer over the more common, and often cheaper, head mounted virtual environment (HMVE)? One of the problems of HMVEs is they have been reported to induce motion sickness. In the CAVE environment, while motion sickness is still possible, these effects are reportedly extremely low (Cruz-Neira, Sanden, and DeFanti 1993). This is due in part to the fact that CAVE images do not need to be updated for every turn of the head or movement of the users body. It can remain static until the user activates an object or indicates that a change of perspective or location is desired. The CAVE environment also permits the user to view their own body, and the bodies of other who are in the environment with them (Leigh, 1996), while they are immersed in the virtual world, which may lessens feelings of claustrophobia or feelings of disembodiment.

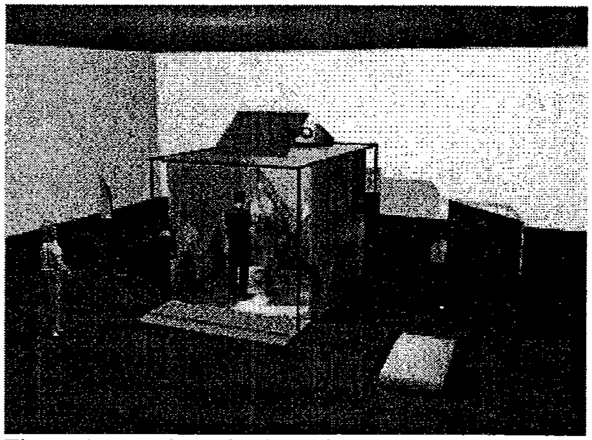

Figure 1: Drawing of CAVE from http://www.cica.indiana.edu/

Why develop a Virtual Tower of Hanoi? According to Fillbrant (1987)"In the psychology of thinking, the Tower of Hanoi is one of the most extensively studied problems. Nevertheless the central question of what and how participants learn during their solving efforts must still be considered as not being answered sufficiently." The Tower of Hanoi (TOH) paradigm and it's isomorphs have found significant performance differences between: right brain vs left brain cognitive styles (Albaile,1996), medicated vs unmedicated ADHD children (Aman 1998), normal hearing vs hearing impaired young adults (Luckner 1992), normal vs obsessive compulsive individuals (Matraix-Coles et al (1999). It has also been used to study the effects of Parkinson's disease (Saint-Cyr 1988). From this sample of research it is clear that the TOH puzzle is a very useful tool in evaluating a wide variety of topics. Thus the flexibility of a virtual TOH might prove itself a useful tool for cognitive research.

How might TOH be related to usability and user evaluation tests? Stanney (1998) commenting on a review of HCI studies states that "The HCI studies indicated that individuals who score low on spatial memory tests generally have longer mean execution times and more first-try-errors." Since TOH can be viewed as a measure of spatial reasoning ability (Albaile, 1996), comparing a users performance to a baseline measure of TOH performance, might be a useful addition to a user assessment battery of tests similar to the Virtual Environment Performance Assessment Battery (VEPAB)(Lampton, 1994).

TOH has been defined as a transformational problem. A transformational problem, is described by Greeno (1978) as, requiring "an initial state and goal state, and permissible operators to transform the initial state into the goal state." From this definition it would seem that researching how people solve transformational problems such as the TOH might provide insights into how people move from an initial state to a goal state when navigating and manipulating modern graphical user interfaces, and virtual environments.

Sweller (1983) discovered that participants who had been trained to solve the 5-ring TOH problem ran into difficulty when asked to then solve a 4-ring TOH puzzle. At first this result seems strange since the 4-ring solution is simpler in that it requires of lower minimum number of moves to reach the solution. However, if participants try to apply the steps of the solution to the 5-disk TOH puzzle to the 4-ring TOH puzzle they end up placing the 4-ring stack on the middle peg instead of on the goal peg where it needs to be. Thus TOH can be employed to evaluate both negative and positive transfer of training issues.

Zhang & Norman (1994) using a TOH isomorphs of coffee cups filled with coffee instead of rings were able to eliminate stacking errors since participants in the experiments instinctively knew not to place the smaller coffee cup on top of a larger coffee cup. In VEs the external representations and the inherent properties of objects can vary significantly from their real world models. Since the objects are not solid they can easily pass through other objects and the user unless this quality is programmed into the object. Since there is no gravity objects will not drop to the ground unless there is a subroutine built into the code that performs this action. They will not make a sound unless this is programmed into their routine, and the sound that is assigned to them is not bound by any natural constraints such as their inherent material composition. And if VE objects were programmed to act as if there was some gravitational force it could be of an infinite degree of strength and could come from a variety of directions. VE objects also cannot be felt with out proper equipment and programs. How might these factors impact performance on the task?

From this introduction it is clear that the potential questions that could be explored with VE's are many. This initial study was designed specifically to evaluate whether there are significant performance differences between the VE and the real world on the TOH task. The dependant variables that were examined in the study are, number of moves to solve the puzzle, and time to move a ring. While no significant performance differences are expected, VE performance may be slower due to the unfamiliar setting and interface. The behavioral differences outlined in the previous paragraph may also be a source of performance difference.

Since objects can be selected with the wand using two methods we looked at whether the click-hold (place wand on object, push and hold button to pick up object, move object, release button to drop object) method for object selection, or the click-release (place wand on object, push and release button to pick up object, move object, push and release button to drop object) method had any impact on performance. It is believed that the click-hold method is a more natural method since the user holds down the button similar to how a person has to hold on to an object in order not to drop it.

If no significant differences are found between the two environments then the virtual TOH can be viewed as a valid tool for studying how various object and environmental manipulations impact task performance. If significant differences are found between the two environments then further research should try to establish the source of the performance differences.

METHOD

Twelve graduate students at Indiana University were used in this study. Their average age was 31.25. All but one of the participants were right-handed. The participants passed tests for color blindness and stereoscopic vision. Only 3 subjects reported that they were familiar with the Tower of Hanoi puzzle. Seven of the subjects reported that they had experienced a virtual environment before, however it was generally not a Cave environment (3D games, full motion simulator, HMD, gloves, one subject had seen a CAVE tour but hadn't used the technology). None of the participants received any form of monetary compensation for their participation. Interest in the CAVE and virtual environments was a sufficient motivating factor.

Three of the real rings were borrowed from a Fischer Price "Rock-A-Stack" ring-stacking toy. The largest ring and the three-pronged base of the puzzle came from a Discovery Toys ring toss game. The virtual rings attempted to match the color to the appropriate size of the real rings. However, during the development of the virtual rings it was discovered that the virtual rings needed to be larger than the size of the real rings due to trouble picking up the rings with the VE wand if ring size had been matched.

Participants in the VE had no tactile feedback so a visual cue, (ring turned red) was given to indicate when an object had been touched. An auditory cue was also provided when the object was picked up (a sliding sound) and released (a dropping kerplunk type sound). The VE also lacked poles to place the rings on but instead had flat black target ring marks on the virtual tabletop to indicate the three valid placement locations. The VE rings also remained suspended in the place where they were released, there was no virtual gravity acting on the objects. Clearly these are potentially significant differences between the two environments. However since these factors may be common in many virtual environments it is worth studying whether their effect on performance is significant.

In an attempt to control for learning effects, six of the participants performed the real world task first and the virtual task second. The other six participants performed the tasks in the reverse order; virtual first, real second. The task was performed with 4 disks. The object of the puzzle is to move the stack of rings from the far left position to the far right position. Movement of the rings must follow the rules that only one ring can be moved at a time, and a ring can only be placed on top of a larger ring or an empty pole position. Since in the VE there was nothing to prevent pulling rings from the bottom of the stack a further rule was added that rings must be selected from the top of a stack. Also in order to test whether wand use affected performance, half of the subjects were instructed to use the click-hold wand method and half were instructed to use the click-release method. Participants were also questioned about their CAVE expectations and experience.

RESULTS

A one-way analysis of variance was performed to compare the performance on the task between the two environments for each of the four trials. The influence of the two wand methods on performance in the VE proved non-significant. While the number of moves to solve the puzzle varied widely (15-113) there was no significant difference detected between the environments. However, the time to move a ring was significantly different between two environments across all four trials, with performance in the CAVE being significantly slower. (Trial 1: $F = 26.557$, $p = 0.000$, $df = 1$; Trial 2: $F=16.057$, $p=0.002$, $df=1$; Trial 3: $F=15.072$, $p=0.003$, $df = 1$; Trial 4: $F = 10.684$, p:0.008, $df = 1$).

ANOVA:	SS	DF	MS	F	P
Trial 1:	22.411	1	22.411	26.557	0.000
Error	8.439	10	0.844		
Trial 2:	15.402	1	15.402	16.057	0.002
Error	9.592	10	0.959		
Trial 3:	3.873	1	3.873	15.072	0.003
Error	2.569	10	0.257		
Trial 4:	9.668	1	9.668	10.684	0.008
Error	9.049	10	0.905		

Table 1. Analysis of variance between the real and virtual environments on time to move an object for trials 1-4. The results indicate that performance was significantly slower in the virtual environment.

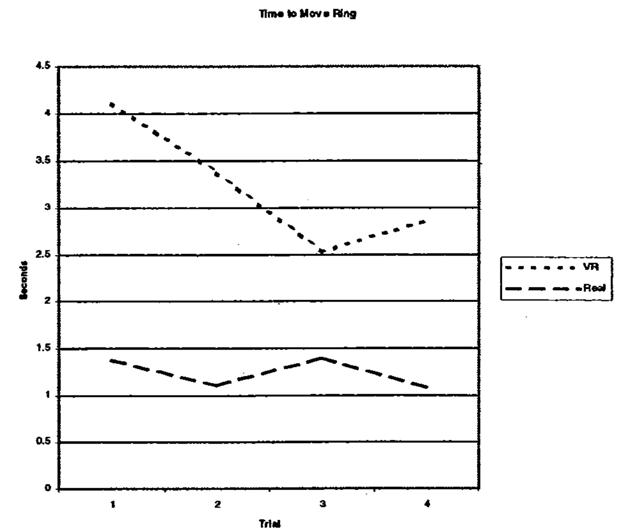

Figure 1: Time to move rings in seconds by environment.

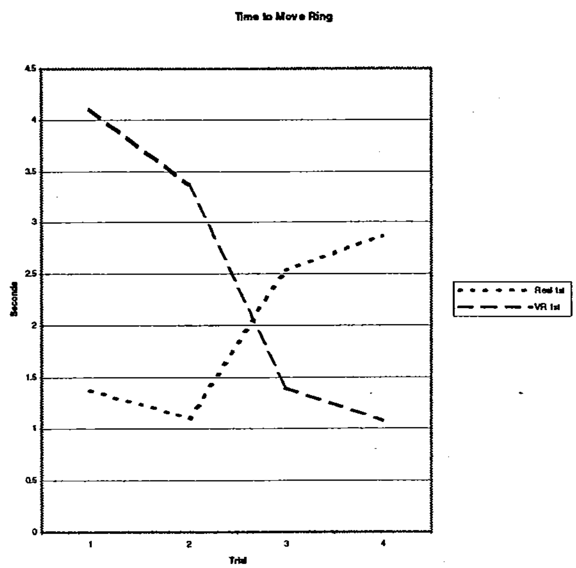

Figure 2: Time to move objects by group. It is interesting to note that while the real world trials appear to have had a beneficial effect on the virtual environment the virtual environment shows no beneficial transfer to the real environment.

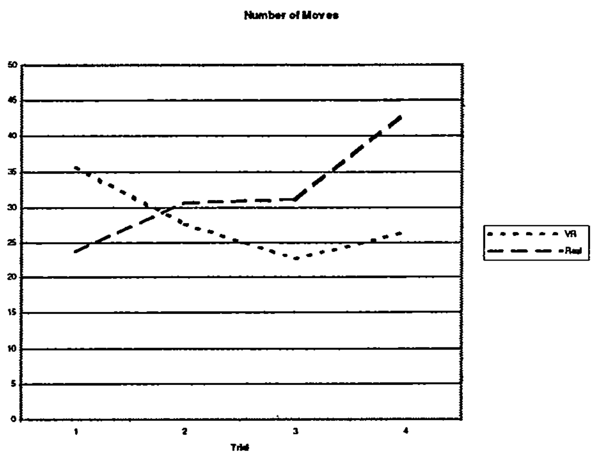

Figure 3: Number of Moves. Analysis indicated no significant difference between number of moves to solve puzzled. Still the increase in number of moves in the real environment is quite unexpected. However it should be noted that there is a single outlier of 113 moves on trial 4 for the real environment.

COMMENTS AND OBSERVATIONS

Participants were asked what they expected the CAVE environment would be like. A few expected that it would be similar to the Holodeck of the television show Star Trek, or the movie "the Matrix". Others had their doubts, stating that they expected it wouldn't quite live up to reality. They expected that the objects would be "boxy", and the motion "choppy". Some felt that performing tasks in the CAVE would be more difficult since they would not be able to perform them exactly the same as they would in the real world.

While observing the participants in the CAVE several recurring problems were noted. Six of the participants had trouble selecting the largest ring. Four subjects put a ring through the virtual table. Four subjects had notable difficulty selecting the rings. Four subjects were observed accidentally embedding rings inside each other. Occasionally a pair of adjacent rings would highlight when the participant attempted to select one of the rings in the pair. Three subjects were noted using moves to adjust ring positions. The line where the floor meets the front wall was visible through the table. A couple subjects broke the rules and attempted to select rings from the middle of the stack instead of from the top. (When this occurred participant were instructed to restore the ring to its former position then continue.)

After experiencing the CAVE the participants were asked to describe their experience. One commented that they felt that their mind was going faster than the system was able to render. They also missed having tactile feedback from the objects. However, one participant stated that they had used the auditory cues as a substitute for the lack of tactile feedback. Some felt that the visuals were rougher than they expected, the objects seemed fuzzy, and not as realistic. Others were surprised at the intensity and realism of the simulation. Several commented that the wand was sometimes difficult to use in selecting the objects, suggesting that gloves may be more natural. Several complained about the lack of gravity and one participant felt the lack of gravity caused them to adjust disk positions in order to match their conceptual model of how the rings should look, and another was annoyed that the rings did not sit on top of each other. One participant commented that they felt the lack of gravity interfered with their ability to remember how they had solved the puzzle in the real world. Others commented that they disliked the fact that they could break normal physical laws such as putting rings through the table.

Even with these comments about the CAVE most subjects felt that the task seemed basically the same in both environments. Strangely even with all the complaints listed above several participants stated that they felt the task was easier in the CAVE! One stated that the VE puzzle seemed more like a game while the RE puzzle seemed more like a test. Perhaps part of this is because the video camera was visible to the participants while they were performing the RE task. It was noted by one of the subject that the lack of objects surrounding the rings in the VE made it easier to concentrate.

While three of the click-release subjects reported no trouble using that wand method, two felt it was counter intuitive, not mouse like, and hard to remember. Participants using the click-hold method stated that they felt it was similar to the mouse, more natural, intuitive, and more like holding objects in real world.

Even the RE task was not free of troubles. During testing it was discovered that the smallest ring could fit within the hole of the largest ring, which occasionally made the smallest ring difficult to pick up. Also, some participants mentioned that the green and blue rings seemed to be more similar in size, and thus harder to distinguish between, than the yellow and orange rings. It was also observed that the poles did on occasion interfere with ring placement. If the ring was not placed correctly over the pole when the ring was dropped it could bounce off the pole causing the participant to break their fluid movement to reclaim the poorly dropped ring.

DISCUSSION

Given the dissimilarities of the environments and the observed difficulties, it is not surprising that there were performance differences between the two environments. It is not clear what the exact source is of this difference. However, it is quite likely due to those factors mentioned earlier in the paper.

Subjects performing the task in the RE often dropped the rings once the rings were over the target pole, however the VE objects did not have gravity to make them fall. This lack of gravity in the VE may have caused the participants to take more time in placing the objects. As mentioned before the VE had no poles. In the RE the poles kept the rings in a neat stack and were utilized with gravity to directing the ring to its proper position. In the VE the participant needed to pay closer attention to where exactly they placed the ring over the target location. The virtual objects were not solid objects and as such could be put through the table or into another object, which could create difficulties when attempting to move the rings.

As mentioned before the virtual rings were larger than the real rings in order to facilitate picking with the wand. However, increased size of the rings should have aided selection of the rings and shouldn't have necessarily slowed movement of the objects. The increase in size did not place the target ring positions outside of the comfortable reach of the participants.

It is interesting to note that there is a decrease in significance across the trials. Thus it may be possible that the difference is due in part to the lack of familiarity in using the wand to move objects in a CAVE environment. Future research may want to compare the performance differences of a more natural interface device such as pinch gloves, or haptic feedback gloves, in comparison to the wand device. One thing that is certain is that visual similarity of objects is not sufficient to ensure similar interaction behaviors between real virtual environments. Other elements such as gravity, object solidity, and tactile feedback without a doubt play a significant role in how we interpret and react to the world around us and should be modeled if a realistic virtual world is the goal.

Reference List

1. Albaile, M. A. (1996). Inferred Hemispheric Style and Problem Solving Performance. Perceptual and Motor Skills, 83, 427-43.

2. Aman, C. J., Roberts, R. J. Jr., & Pennington, B. F. (1998). A Neuropsychological Examination of the Underlying Deficit in Attention Deficit Hyperactivity Disorder: Frontal Lobe Versus Right Parietal Lobe Theories. Developmental Psychology, 34(5), 956-969.

3. Burdea, G., Richard, P., & Coiffet, P. (1996). Multimodal Virtual Reality: Input-Output Devices, System Integration, and Human Factors. International Journal of Human Computer Interaction, 8(1), 5-24.

4. Cruz-Neira, C., Leigh, J., Papka, M., Barnes, C., Cohen, S. D., Engelman, R., Hudson, R., Roy, T., Siegel, L., Vasilakis, C., DeFanti, T. A., & Sandin, D. J. (1993). Scientists in Wonderland: A Report on Visualization Applications in the CAVE Virtual Reality Environment. In IEEE Symposium on Research Frontiers in Virtual Reality (pp. 59-66). Los Alamitos, California: IEEE Computer Society Press.

5. Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A. (1993). Surround Screen Projection Based Virtual Reality: The Design and Implementation of the CAVE. In SigGraph 93 (pp. 135-142).

6. Greeno, J. G. (1978). Natures of problem solving abilities. In W. K. Estes (Ed.), Handbook of learning and cognitive processes (Vol. 5). Hillsdale, N. J.: Erlbaum.

7. Heim, M. (1993). The Erotic Ontology of Cyberspace. M. HeimThe Metphysics of Virtual Reality (pp. 83-108). Oxford: Oxford University Press.

8. Heim, M. (1993). The Essence of VR. M. HeimThe Metphysics of Virtual Reality (pp. 109-138). Oxford: Oxford University Press.

9. Lampton, D. R., Knerr, B. W., Goldberg, S. L., Bliss, J. P., Moshell, J. M., & Blau, B. S. (1994). The virtual environment performance assessment battery (VEPAB): Development and evaluation. Presence: Teleoperators and Virtual Environments, 3(2), 145-157.

10. Larijani, L. C. (1994). Applications for Science & Engineering. In L. C. LarijaniThe Virtual Reality Primer (pp. 165-178). New York: McGraw Hill, Inc.

11. Leigh, J., & Johnson, A. In Proceedings of the International Conference on Multimedia Computing and Systems (pp. 20-23). Los Alamitos, Calif.: IEEE Computer Society Press.

12. Luckner, J. (1992). Problem Solving: A Comparison of Hearing Impaired and Hearing Individuals. Journal of the American Deafness and Rehabilitation Association, 25(4), 21-27.

13. Stanney, K. M., Mourant, R. R., & Kennedy, R. (1998). Human Factors Issues in Virtual Environments: A Review of the Literature. Presence, 7(4), 327-351.

14. Zhang, J., & Norman, D. A. (1994). Representations in Distributed Cognitive Tasks. Cognitive Science, 18, 87-12.

I would like to thank Dr. Katy Boerner for programming the environment. And Dr. Andrew Dillon for advising me on the project.