Digital Dome Versus Desktop Computer in a Learning Game for Religious Architecture

Jeffrey Jacobson
PublicVR

Please address all correspondence to:
Jeffrey Jacobson, Ph.D.
Director, PublicVR
333 Lamartine St.
Jamaica Plain, MA 02130
617-435-0517
jeff@planetjeff.net
www.publicvr.org

Paper presented at the
Annual Meeting of the American Educational Research Association
Denver, CO, April 30 – May 4, 2009

Digital Dome Versus Desktop Computer in a Learning Game for Religious Architecture

Jeffrey Jacobson
PublicVR

An immersive visual display can help the student when concepts are anchored or embodied in an information-rich space and when an egocentric viewpoint is revealing. For example, ancient Egyptian temples reveal the entire culture by juxtaposing language and symbol in architecture meant to be understood from the inside. Students, aged 11 through14, played an educational computer game based on a (virtual) temple. Those using a digital dome had better factual recall (P < 0.05) and conceptual learning (P < 0.001) than those using a desktop computer. The difference was not evident in a structured written test but apparent in student-made informal video tours of the virtual temple. Students with low visual reasoning skills appeared to benefit most in conceptual learning (P < 0.01).

Introduction

We present empirical evidence that visually immersive displays can improve learning in educational virtual reality applications. For example, the digital dome in Figure 1 immerses the audience in the virtual environment, making the live tour more effective. Our long-term goal is to understand why, how and *when* Visually Immersive Virtual Reality (VIVR) is useful for learning and what its best use is in a larger curriculum. As with all other media, VR does not replace any vehicle but adds to the tools available to the educator.

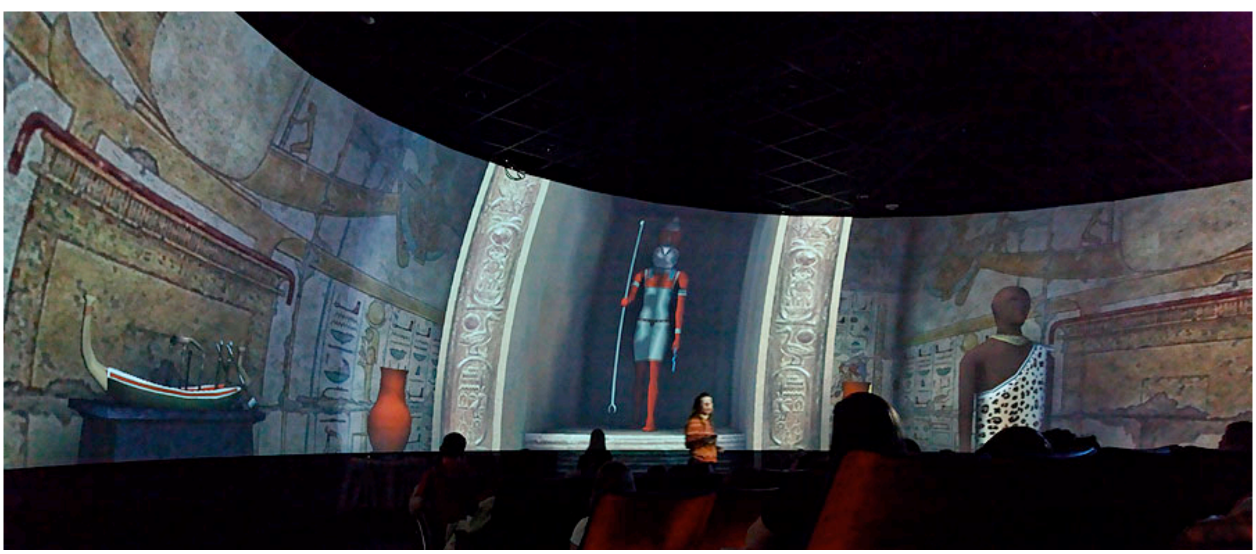

FIGURE 1: *The Virtual Egyptian Temple in the Earth Theater*

The value of VIVR for acquiring training is well-established (Ellis, 1991), but research on its usefulness for traditional education topics (e.g., science) has been problematic. Despite VIVR's apparent advantages for education that gave rise to hundreds of educational projects, we could find only fourteen formal experimental studies (Moreno, 2002; Winn, 1997, 2001; Jackson, 2000; Bowman, 1999; Dede, 1999; Roussos, 1999; Salzman, 1998, 1999; Osberg, 1997; Byrne, 1996; Rose, 1996; Jacobson, 2008; Bailenson, 2008; Limniou 2007) and one using activity theory (Roussou, 2007). Interactive visually-immersive displays were cost-prohibitive until recently (Lewis, 2002), they are still complex to employ, and evaluation of educational VIVR is tricky. Nevertheless, clear trends emerge.

VIVR works best when the student has an *egocentric* viewpoint in the virtual environment, meaning he or she is (virtually) inside something interesting (Dede, 1999). The view surrounds the viewer with information, making important relationships easier to read. Examples are a magnetic field (Salzman, 1999) or a simulation of Puget Sound with added visual representations of items, water temperature, flow, and salinity (Winn, 2001).

This study was part of a larger collaboration of PublicVR, the Carnegie Museum of Natural History in Pittsburgh (CMNH, 2009), and the School of Information Sciences at the University of Pittsburgh (PITT, 2009). The author was the principal investigator, and the list of other contributors is too long to list here. See Jacobson (2008).

Protocol

Our educational computer game is based on our Virtual Egyptian Temple (Jacobson, 2005, 2009; PublicVR, 2009). In the game, the student selects a feature of the temple; then the virtual priest explains what it is, what it means, and how it fits into the whole. The student advances from one area of the temple to the next by successfully answering the priest's questions for each area, and the player "wins" by answering all questions for the inner sanctuary.

FIGURE 2: *A student plays Gates of Horus in the Earth Theater*

We conducted all testing at the museum's Earth Theater (CMNH, 2009) with middle school students, ages 11 through 14. They were old enough to handle the material and the test instruments, but young enough to be enthusiastic about planned activities. They are also familiar with computer games and activity-oriented learning from their earlier schooling

When the students arrived, we pre-tested their visual reasoning skills using Raven's Progressive Matrices (Raven, 1961). Then each student was randomly assigned to the control group or treatment group. 40 students played *Gates of Horus* on a desktop computer (the control group) while 27 students each used the Earth Theater's digital dome as a personal display (the treatment group). Finally, each student produced a short documentary video describing the temple in an unstructured 15-minute show-and-tell – the Video Test (Figure 3). There, the student navigated the Virtual Egyptian Temple, which was displayed in a wall projection. We did take other measurements as part of a larger study (Jacobson, 2008), which is beyond the scope of this paper.

FIGURE 3: *A student gives a tour of the temple*

Experimental Hypothesis

> *Students using the visually immersive display while playing the educational computer game will demonstrate more knowledge, immediately afterward, than those using a standard desktop monitor.*

Analysis

We evaluated each student-made video by having a grader (rater) fill out our questionnaire, which lists the facts we expected the video to include (Figure 4). The grader could assign full or half credit for how the student described each fact. All student videos were evaluated three times, once by each of three graders. Our goal was to aggregate the three raters' scores to produce a single set of grades, which is more reliable than the judgments of a single person.

However, we combined only sets of scores that agreed sufficiently for the aggregate to be meaningful. We used the Fleiss' Kappa test (Fleiss, 1981) for interrater reliability analysis, which showed grades for 45 out of the 75 facts (60%) having acceptable levels of rater agreement ($P < 0.5$). Most of the rejected grades involved facts few students mentioned, leaving too little information to determine rater agreement. Fortunately, accepted scores accounted for 89% of all points awarded, quite sufficient for comparing performance between groups. Within the accepted data, we were able to average raters' data sets together for an aggregate set of scores, which we used in further analysis.

Next, we compared students' average scores for all factual learning in the treatment group to the averages in the control group, using a standard T-Test. The treatment group showed a significant advantage of **P=0.0458** in factual learning. This is the main result of our study.

To better understand the main result, we made the same comparison for students' scores across each of the 45 individual facts on which they were tested. For each fact, we compared the aggregate scores for the treatment and control, using the nonparametric Mann-Whitney Test. The standard T-Test would not have been appropriate, because we allowed only distinct grades (none, half, full) so our data is not continuous. Another reason is that the data are highly skewed, with the "none" grades greatly outnumbering the other two.

Non-parametric tests are less sensitive than those which work with continuous functions (i.e. ANOVA, T-Test, etc.), but we did get some useful detailed results. The treatment group did better, at $P < 0.05$ for nine facts and at the $P < 0.10$ for another five. The effect was stronger when we looked at results for the sanctuary, which is the last and most important area of the temple, where all its themes come together. Students did better at the $P < 0.05$ level for four of the 13 facts attached to the Sanctuary and at the $P < 0.10$ level for three more. There were no cases the control group did better than the treatment group.

It is important that these 14 out of 45 facts where the treatment group did provably better does not tell the whole story. For many of the other 31 facts, the treatment group also did better, but at a statistically significant level. However, this was enough for the *main result* based on students' aggregate scores to a significant treatment group advantage ($P < 0.5$). Furthermore, the stronger result Sanctuary ($P < 0.04$) is arguably a stand-in for the main result. Finally, factual learning is not the true strength of VR-based learning, and may well be a second-order effect of deeper understanding.

Our Egyptologist (Holden) evaluated the videos for conceptual learning to complement the other evaluators' relatively rote scoring of factual and lower-level conceptual learning. Students using the immersive display generally did better at explaining the most important concepts ($P < 0.001$). The treatment group also tended to do better with the particular facts most closely linked to those key concepts.

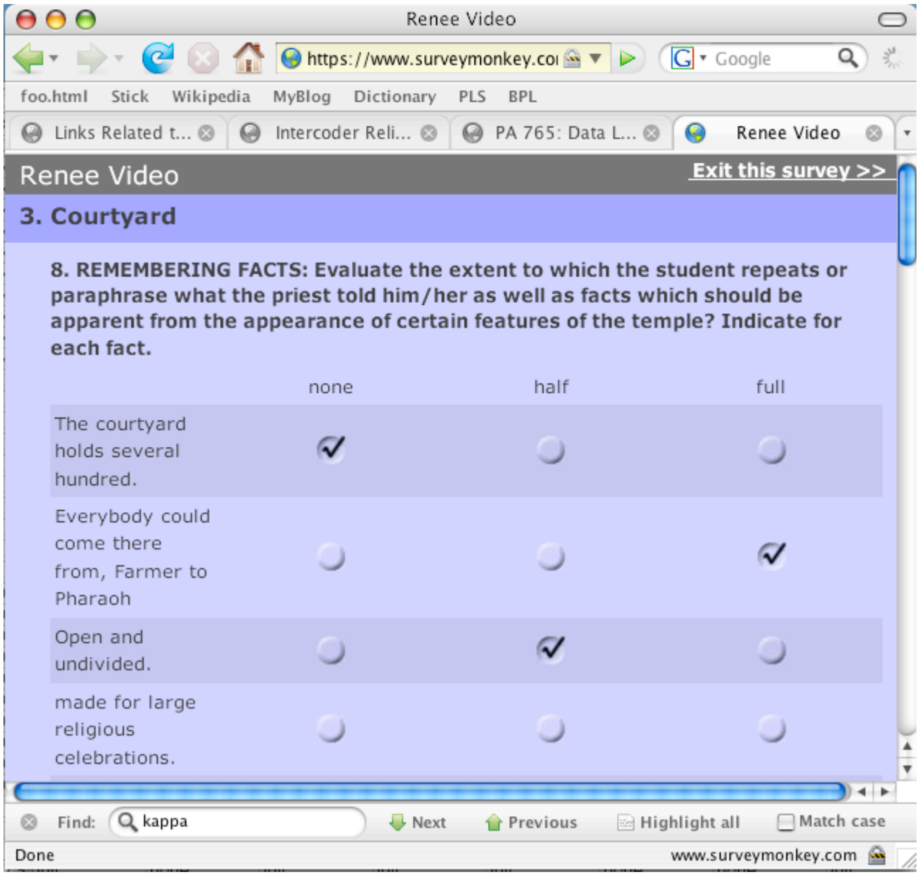

FIGURE 4: *Part of a grader's scoring sheet. The courtyard is part of the temple, and "Renee" is the student's code name.*

Students also completed the Written Test, containing multiple-choice and short-answer questions. *Four* graders evaluated students' short answers in much the same way as graders evaluated the videos. They awarded full or partial credit for stating expected facts, in the results aggregated through interrater reliability analysis. The results demonstrated that *Gates of Horus* is an effective learning tool (Jacobson, 2009). However, further analysis did not prove an advantage for the treatment group, with P=0.67 for average short-answer scores, and similar disappointment from the multiple-choice questions. See Jacobson (2008) for details.

Discussion

We believe that the visually immersed learners developed a more coherent and complete mental map of the temple and the facts and concepts anchored to it. The temple is more easily readable in the immersive display, reducing cognitive load for the user by showing more features and displaying those features in their proper configuration. The immersive display enables the student to search the virtual space more efficiently, with more looking and less navigating.

An alternative explanation is that students in the Earth Theater were simply more engaged because of the sense of *presence,* the feeling of "being there" in the virtual space. Most educational literature for visually immersive VR cites *presence* as the primary factor for enhanced learning. However, every learning experiment depending solely on presence failed to show a learning effect (Moreno, 2002b; Byrne, 1996; Rose, 1996). We agree with Bailenson (2008) that presence is too poorly defined to be isolated as an experimental factor. A clear educational value of visual immersion comes from the way it can reveal information in an egocentric view (Dede, 1999) of a virtual environment. More broadly, the most important educational value of VR in all its forms is its ability to provide an adaptive learning environment (Winn, 2003).

Finally, we see an interaction effect between Holden's evaluation of understanding and the data from Raven's Progressive Matrices (RPM). It appears that students with low RPM scores benefited more from the immersive display than those with high RPM scores (P=0.01). The result parallels a study by Winn (1997) in which students who made poor grades benefited more from using an immersive display in Winn's educational VR application (Winn, 2001). This interaction effect was not evident in scoring for factual data.

We believe that the written test was not sensitive to genuine learning difference between the two groups of students, at least for our sample size. Multiple-choice and short-answer tests provide much information to the test-taker in the form of questions, providing a framework for the information the student is expected to produce. This downplays differences in students' ability to mentally organize and structure the information on their own, which is more important than factual learning.

By contrast, the video test is unstructured, providing only the virtual temple itself as a scaffold for recall and presentation. (The written test provides no imagery.) The video test is also more like real life. One's understanding of an information-rich space (e.g., religious architecture) is most likely to be used in the space itself. Overall, students who played *Gates of Horus* in the immersive display recited more facts during their video test than those who used a standard desktop, as we described in the results section.

Not only can educational VR help all students learn certain topics, it may reach certain students who are underserved in the current educational system. Our learning game is entirely visual and verbal, while the video test has no written component, so both treatment and test bypass literacy. Many smart children do poorly on written tests for a variety of reasons, and the use of visual media is a way to reach them. In our future research, we would like to explore where and when Visually Immersive VR is helpful to all students and when it helps students who are currently underserved.

REFERENCE LIST

Bailenson, J. N., Blascovich, J., Beall, C., Lundblad, N, Jin, M. (2008). The Use of Immersive Virtual Reality in the Learning Sciences: Digital Transformations of Teachers, Students, and Social Context. The Journal of the Learning Sciences, 17, 102-141 ISSN: 1050-8406 print / 1532-7809 online.

Bowman, D. A., Hodges L. F., Allison, D., Wineman, J. (1999). The Educational Value of an Information-Rich Virtual Environment. Presence Teleoperators and Virtual Environments, 8(3), 317-331.

CMNH (2009). Carnegie Museum of Natural History, 4400 Forbes Ave., Pittsburgh, PA 15213. http://www.carnegiemnh.org/.

Dede, C., Salzman, M. C., Loftin, R. B., Sprague, D. (1999). Multisensory Immersion as a Modeling Environment for Learning Complex Scientific Concepts. In Computer Modeling and Simulation in Science Education. Springer-Verlag.

Ellis, S. R. E. (1991). Pictorial Communication in Virtual and Real Environments (2nd ed.). London: Taylor and Francis.

Fleiss, J. L., Levin, B., Cho, M. (1981). Statistical Methods for Rates and Proportions. ISBN: 0-471-52629-0.

Jackson, R. L., Winn, W. (1999, Dec. 12-15). Collaboration and Learning in Immersive Virtual Environments. Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference, Stanford University, Palo Alto, California.

Jacobson, J., Holden, L. (2005). The Virtual Egyptian Temple. World Conference on Educational Media, Hypermedia & Telecommunications (ED-MEDIA), Montreal, Canada.

Jacobson, J. (2008). Ancient Architecture in Virtual Reality; Does Visual Immersion Really Aid Learning? Dissertation, School of Information Sciences, University of Pittsburgh. http://planetjeff.net/Jacobson2008.pdf.

Jacobson J. (2008). Gates of Horus. http://publicvr.org/html/pro_gates.html.

Lewis, M., Jacobson, J. (2002). Game Engines in Scientific Research. Communications of the ACM, 45, 27-31.

Limniou, M., Roberts, D., Papadopoulos, N. (2007) Full Immersive Virtual Environment CAVE™ in Chemistry Education. Computers and Education, 51 (2008), 584-593

Moreno, R., Mayer, R. E., Lester, J. C. (2000a). Life-Like Pedagogical Agents in Constructivist Multimedia Environments: Cognitive Consequences of their Interaction. ED-MEDIA 2000.

Osberg, K. M. (1997a). Constructivism in Practice: The Case for Meaning-Making in the Virtual World. Unpublished doctoral dissertation, University of Washington, Seattle, Washington.

PITT (2009). University of Pittsburgh, School of Information Sciences, 135 North Bellefield Avenue, Pittsburgh, PA 15260. http://www.ischool.pitt.edu.

PublicVR (2009). http://publicvr.org.

Raven, J. (1958). Standard Progressive Matrices. San Antonio, TX: The Psychological Corporation.

Rose, H. (1996). Design and Construction of a Virtual Environment for Japanese Language Instruction. Unpublished Master's thesis, University of Washington.

Roussos, M., Johnson, A. Moher, T., Leigh, J., Vasilakis C., Barnes, C. (1999). Learning and Building Together in an Immersive Virtual World. Presence, 8(3), 247-263.

Roussou, M., Oliver, M., Slater, M. (2007). Exploring Activity Theory as a Tool for Evaluating Interactivity and Learning in Virtual Environments for Children. Journal of Cognition, Technology & Work, Springer, ISSN: 1435-5558 print / 1435-5566 online.

Salzman, M. C., Dede, C., Loftin, R. B., Ash, K. (1998). Using VR's Frames of Reference in Mastering Abstract Information. Third International Conference on Learning Sciences, Charlottesville, VA.

Salzman, M. C., Dede, C., Loftin, R. B., Chen, J. (1999). A Model for Understanding How Virtual Reality Aids Complex Conceptual Learning. Presence: Teleoperators and Virtual Environments, 8(3), 293 - 316.

Spring, M. (1992). Virtual Reality and Abstract Data: Virtualizing Information. Virtual Reality World, 1(1).

Winn, W., Windschitl, M., Hedly, N., Postner, L. (2001). Learning Science in an Immersive Virtual Environment. Presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Winn, W. (2003). Learning in Artificial Environments: Embodiment, Embeddedness and Dynamic Adaptation. Tech., Inst., Cognition and Learning, 1, 87-114.

Winn, W. (1997). The Impact of Three-Dimensional Immersive Virtual Environments on Modern Pedagogy (Technical Report R-97-15). Seattle: Human Interface Technology Lab, http://www.hitl.washington.edu/publications/r-97-15/.

Zeltzer, D. (1992). Autonomy, Interaction and Presence. Presence: Teleoperators & Virtual Environments, 1(1), p 127-232.