Ancient Architecture in Virtual Reality;

Does Visual Immersion Really Aid Learning?

Jeffrey Jacobson, PhD

University of Pittsburgh, 2008

The following document is the summary section of my full dissertation, which you can see at

http://PublicVR/Jacobson2008.pdf

7.0 SUMMARY AND CONCLUSION

This study explored whether students benefit from an immersive panoramic display while studying subject matter which is visually complex and information-rich. Specifically, middle-school students learned about ancient Egyptian art and society using an educational learning game, *Gates of Horus*, which is based on a simplified virtual model of an Egyptian temple. First, we demonstrated that the game is an effective learning tool by comparing written post-test results from students who played the game and from students in a no-treatment control group.

Figure 46. Virtual Egyptian Temple and the High Priest

Next, we compared learning results of two groups of students who had the same mechanical controls, but one group saw the temple in a visually immersive display (a partial dome) while the others saw it on a standard desktop monitor. The difference appeared when each student gave a verbal show-and-tell presentation of the Temple and the concepts and facts related to it. During that exercise, the student had no cognitive scaffolding other than the virtual temple, on a small wall projection, which the student navigated during the presentation. The other major tests were questionnaires, which by their nature provide a great deal of scaffolding for the task of recalling the required information. For these tests we believe that this scaffolding aided students' recall to the point where it overwhelmed the differences produced by any difference in the display.

We conclude that the immersive display provides better supports for the student's learning activities for this material. To our knowledge, this is the first formal study to show concrete evidence that visual immersion can improve learning for a non-science topic.

This section summarizes the entire dissertation and contains many links to the main body of the text. We recommend that the reader understand this summary first, and refer to other areas of the dissertation for more detail.

7.1 BACKGROUND

Section 2.6, p62 catalogs previous educational research in Immersive Virtual Reality and presents its theoretical underpinnings. Here, we summarize what we learned from that survey and what we set out to accomplish.

Looking for previous studies in the educational use of Immersive Virtual Reality, we find only *twelve* (Rose, 1996; Byrne, 1996; Winn, 1997; Osberg, 1997a; Salzman 1998; Salzman, 1999; Roussos, 1999; Bowman, 1999; Dede, 1999; Jackson, 2000; Winn, 2001; Moreno, 2002a) which used a formal experimental design (Campbell, 1963), and not many more studies which use other formal methods (Roussou, 2006, 2008). We attribute this small number to the historically high cost of the necessary technology, which limited most development in the 1990s to the computer scientists and most applications to the military. Nevertheless, there is still much interest in VR in the educational community, as evidenced by the very large number of Desktop VR applications and learning experiments (Cobb, 2002). Fortunately, the cost of ImmersiveVR is decreasing (Lewis, 2002; Dalgarno, 2002a; Young, 2000; Pivec, 2003b; Bruckman, 2002a; DeLeon, 2000; Tougaw, 2003; Stang, 2003; Jacobson 2005i; CaveUT, 2008), which enabled us to do our study. We believe that this decrease in cost is already leading to more educational research in immersive virtual reality.

7.1.1 What is Virtual Reality?

The meaning of the term "Virtual Reality" (VR) continues to evolve and differentiate, as disparate research communities find new ways to use it. We present our own working definition, which we state as criteria:

- **A Three-dimensional Space:** The user perceives an illusory three-dimensional "space" defined by the (computer graphic) objects within it and known as the *Virtual Environment* (VE). This illusion is composed of information rendered into a form the user can perceive—usually through computer graphics artwork.

- **Interaction:** The user must be able to interact with the virtual environment in a meaningful way (Zeltzer, 1992). Virtual objects and actors respond to the user in a manner consistent with their purpose and the overall theme of the environment. For example, a virtual cat in an Egyptian temple might come near the user, but avoid being "touched."
- **Autonomy:** Many objects or processes in the virtual environment will (appear to) continue to operate without input from the user (Zeltzer, 1992). For example, a virtual dog digging holes in a virtual garden would continue to produce more holes, whether or not the user was nearby.
- **(Thematic) Presence:** The user must have some imaginary location *and* identity within the virtual environment. An actual virtual body and a co-located viewpoint may be clearly represented or merely implied. The important thing is that the user has a location in the virtual space and a role in the narrative.

For practitioners in the education and virtual heritage communities, these criteria are enough for an application to be considered Virtual Reality. We will call them examples of ***Desktop VR,*** because they nearly always use a standard computer monitor and keyboard and mouse. However, traditional VR researchers and computer scientists have a further requirement:

- (**Sensory) Presence:** The interface informs the senses so that the user sensorially feels like s/he is at a particular location in the virtual environment (Zeltzer, 1992). Perspective correction and other aspects of a visual display define the user's *egocenter,* which is the user's location in the virtual environment (Psotka, 1996). At a bare minimum, the display must produce a very wide view for the user, as with a digital partial-dome theater (e.g. a Planetarium) or a *Head Mounted Display* (HMD). See section 2.3, p17, for examples.

In the following discussion, we will call applications satisfying all five criteria, ***Immersive Virtual Reality*** (Immersive VR) and those which satisfy only the first four criteria, ***Desktop Virtual Reality*** (Desktop VR). We retain ***Virtual Reality*** (VR) as a general term referring to both.

7.1.2 VR Is Expressive

Virtual Reality provides the educator with new ways to represent objects and systems more effectively than with other media. In VR, the user can interact with simulations of things that could be perceived in the real world (Roussos, 1999), because they are too small, too big, no longer exist, do not exist yet, dangerous, far away, or simply inconvenient. Students can make mistakes in a virtual environment safely and cheaply, which allows for learning activities not possible in the real world. A virtual environment could simulate dynamic systems such as ocean currents planetary motions, changes in electrostatic fields, or social behavior in a troop of gorillas. These simulations become especially powerful teaching tools when the student can participate in them, giving the student an inside view (egocentric) and the ability to experiment with the system (Winn, 1999; Bowman, 1999; Dede, 1999)

VR is also a powerful means of communication. Students and educators can collaborate over great distances using a networked virtual environment as a collaborative space (Cobb, 2002; Dede, 2004; Bruckman, 2002a; Andrews, 2002; Räihä, 1997; Raalte, 2003; Santos, 2002). The multisensory interface and potential methods of interaction allow VR applications to communicate a lot of information to and from the user (Bowman, 2002; Mayer 2001c). VR creates significant opportunities for non-written and even non-verbal communication, which can be very useful for certain situations and students. Autonomous agents can represent people conducting their business or interacting with the user (Ulicny, 2002). Pedagogical agents in the environment can guide and facilitate learning (Economou, 2001).

Nevertheless, virtual reality is still just one more form of media with comparative advantages and appropriate uses. VR should not be used as a general replacement for anything, but instead as part of a larger curriculum (Bowman, 1999; Wickens, 1992; Hay, 2000).

7.1.3 VR Can Accommodate the Learner

The theoretical underpinnings of almost all educational experiments in Immersive VR, as stated by the authors, center on the idea of learning as active, self-directed, and context-dependent. The student does or makes something, alone or in a group to gain new knowledge. The student must *construct* new personal knowledge by understanding the lesson in terms of his or her prior knowledge (Bloom, 1956) and perception of the world (Winn, 2003b). The instructor helps the students learn how to learn, facilitating and focusing the students' own process of exploration.

This model of learning is called *Constructivism* (Jonassen, 2000c). In *Activity Theory*, this is called the "co-construction of shared knowledge" (Vygotsky, 1978; Jonassen, 2000b; Jackson 2000). Hay (2000) calls it "investigation-based learning." Many of the Immersive VR learning studies specifically claim a *Constructivist* basis (Fallman, 1999; Dalgarno, 2001b, 2002a, 2002b). The basic idea is that VR gives the instructor many ways to produce (and control) instructional situations for the students to work through. These special-purpose virtual environments can be rich in meaning and complexity or elegant and focused, as needed.

A very important part of this process is ***Conceptual Change*** (Gagne, 1987). This is the moment when students must abandon some misconception when confronted with both counter-evidence and a better explanation of what they thought they knew. Many Immersive VR learning studies were designed to produce this effect (Dede, 1999; Moher, 1999; Winn, 2001; Windshitl, 2000; Jackson, 2000; Johnson, 1999a). Conceptual Change is especially helpful in teaching science topics, where students often have misconceptions of natural processes.

Other researchers sought to more clearly define the learning process, adding approaches from *Cognitive Theory* (Winn, 2003a; Moreno, 2002b; Mayer, 2000b). We prefer Winn's (2003a) conditions for optimal learning with VR. The student must be **embedded** in the virtual world, meaning that learning tasks and overall experience are defined in whole or in part by the virtual reality application. The student must also be **embodied**, in that the learner's physical body is an essential part of the process. This could be as simple as requiring the student to perform physical actions, or as subtle as engaging the student's senses in a particular way. Finally, the student must engage in **dynamic adaptation,** always adjusting to changing circumstances. The individual and his or her environment evolve together, responding to each other. We use a fourth aspect, **Connectedness,** where the learner is connected to information, virtual entities or processes, other students, instructors, or other users. In our own research, we strive to develop learning experiences which have these properties to the extent possible.

We restate that researchers in educational Virtual Reality believe learning is an active and individual process, which good curriculum supports. Virtual Reality can support complex and meaningful interaction using metaphors which are naturalistic and easy to understand. The virtual environment can respond to a student based on where that student is in the learning process.

7.1.4 Sensory Presence Is Not Enough

We believe that sensory presence is one of many desirable forms of engagement. For example, an interesting conversation with an artificially intelligent character (an *agent*) may fascinate the student (Charles, 2007). Perhaps the student is very interested in the topic or intent on winning a VR-based learning game. Having a central place in the narrative (thematic presence) is certainly engaging. We believe that different types of engagement are mutually reinforcing and tend to have common causes, such as good application design.

Almost everyone involved in Educational Virtual Reality believes that sensory presence enhances student engagement and therefore facilitates learning. Nearly every article on the subject clearly states or strongly implies this, e.g., Furness (1997). Surveying the literature, we see that a sense of *presence* self-reported by the students is definitely correlated with learning in VR. However, correlation does not imply causation.

No experiment which relied solely on sensory presence to enhance learning was successful in showing any advantage for immersion (Moreno, 2002b; Byrne, 1996; Rose, 1996; Salzman, 1999). In these studies, sensory immersion was *not* relevant to the learning activities—their students could have done those activities equally well with a non-immersive display. In each experiment, there was no significant difference in learning between students who had an immersive display and those who did not. By contrast, all of the studies which did not show an advantage for immersion (Dede, 1999; Salzman, 1998; Winn, 2001) had the student perform learning tasks which depended on that student "being" in a particular location within the virtual environment. This allowed the student to perceive information and interact with the environment in a particular way.

We conclude that the value of an immersive display is *not* determined by the sensory presence it elicits, but by its functional support for the learning activity.

7.1.5 Immersion Must Support the Learning Activity

We are interested in when and how *Immersive* Virtual Reality can be helpful for learning. To be effective, an Immersive VR learning application must (1) be reliable, stable and useful, (2) support clearly defined learning activities, (3) employ some of the capabilities of VR, and (4) use the visual immersion to functionally support learning activities. The first two criteria are basic to any learning software. The third criterion is that the learning activities require some of the capabilities of VR to support a learning experience which other media could not do as well. Generally, one would ask this of any proposal to use a particular interactive medium for a particular learning activity. Similarly, one must use sensory immersion (Immersive VR) only if it is an integral part of a desirable learning activity which is not otherwise available.

Generally speaking, immersive virtual reality is appropriate when learning activity requires interaction between the student and something that is difficult or impossible to simulate or encounter in real life. Immersive VR will be more advantageous the more complex, context-dependent, and physical the task is. This is why pilots start learning to fly with flight simulators, historically, the first examples of Immersive VR technology.

Learning to operate a plane is usually described as *training*, while *education* is usually thought of as classroom learning. The distinction is vague, and we mention it only because there is much literature on using virtual reality for training in a variety of fields. The value of immersive virtual reality for training is not in doubt. We are interested in seeing how effective it can be as a tool for the major subjects taught in schools.

Salzman (1998), Dede (1999), and Winn (1997, 2001) have had success with topics in science, where the student works with an interactive representation of some otherwise invisible process. Many lessons in science are cognitively difficult to absorb, but lend themselves to helpful representations in virtual reality. We believe this is a fruitful area of research and look forward to further developments.

7.1.6 Virtual Heritage

Virtual heritage (VH) is the use of electronic media to recreate or interpret culture and cultural artifacts as they are today or as they might have been in the past (Moltenbrey 2001; Roehl, 1997). We believe this is an excellent topic for VR learning research, given the centrality of visual artifacts and places, and the importance of history in real life and the classroom.

By definition, VH applications employ a three-dimensional representation of something and the means used to display it, from still photos to immersive virtual reality. This is a very active area of research and development, (Michell, 2000; Champion, 2004b; Champion, 2004c; Addison, 2000; Roehl, 1997; Stone, 2002; Levy, 2004) and most of it is intended for educational use. The majority of VH applications are architectural reconstructions, centered on reconstructed buildings or monuments, and most of them use VRML technology. A handful of VH applications illustrate topics on ancient Egypt (Kufu, 2004; Economou, 2001; Lehner, 2003; Michell, 2000; TutTomb, 2001).

These three-dimensional objects are "well-integrated" in the sense that much cultural information is encoded in the way the space looks to an observer. Therefore, applications in virtual heritage have much to gain by using virtual reality. While most VH applications are limited to the desktop (Kameas, 2000), some employ Immersive VR (iGrid, 2000; Pape, 2000; Park, 2003; PublicVR, 2008) some make excellent use of augmented reality (Papagiannakis, 2004a, 2004b). In our study, students interact with the Virtual Egyptian Temple (Jacobson, 2005e) using Immersive VR.

7.1.7 Learning Games

The game paradigm is an excellent way to center interaction on the user in a flexible virtual environment. Educational researchers have been interested in harnessing games as a vehicle for learning for a long time (Avedon, 1972), and today such efforts have attracted significant interest and resources (Squire 2003, 2007; Kirriemuir 2004). Goal-seeking activities are especially effective and can be cast in the form of a game (Champion 2004b; DeLeon, 2000). Every game is based on a microworld of some type, and IVR interfaces are optimal for interacting with many useful types of virtual environments, especially those with a high level of visual fidelity.

Today, computer games are a central activity in popular culture, with millions of children now playing them. Through video games and other electronic media, many students have developed a high degree of video literacy, comfort and competence with fast, information-dense input. Also, the game industry has developed a large quantity of powerful, flexible software that can be adapted for educational use, and much of it has already been adapted to IVR (Squire 2002, Kirriemuir 2004, and Jacobson 2005i). These new developments put many types of educational virtual reality within reach of educators and their institutions (Lewis, 2002). The key advantages of a game-based learning are:

1. The student's intense investment toward reaching a goal defined by the educator or designer.
2. Continual feedback for the student while interacting with the system.
3. A high degree of student involvement or investment in the activity itself.
4. The potential for intense student concentration on the learning task.

We believe that educational games are a special case of adaptive media and are well-suited to Constructivist learning activities *and* to Virtual Reality (Brusilovsky, 2003c).

The key is to make the goals of the game serve the student's learning goals and the broader curriculum. The student benefits little if the game is designed as some kind of a wrapper around the information he or she is expected to learn, because the student might play the game for its own sake and quickly forget the topic matter. For example, it would be easy to design a game in which the student gains points for solving riddles or remembering facts. However, it would be much more effective if the goal of the game was to accomplish something within the context of the topic matter. Winn (2001) provides an excellent example of good conceptual design. In his study, students adjust environmental factors in a simulated part of the world to find an optimal solution to global warming. The activity is definitely a game, although Winn does not describe it as such in his paper or to his subjects. The student simply enjoyed and learned. In our study, we structured the student learning experience as a game.

7.2 EXPERIMENT

7.2.1 Rationale

The goal of our study was to determine whether a visually immersive display can have a provable advantage over a non-immersive display for topics in cultural heritage.

We chose the Egyptian temple as the sample topic for our learning study, because the temple was absolutely central to Egyptian life and culture, *and* because it is appropriate for the virtual reality medium. The temple itself, the hieroglyphics, the painted images, and the conduct of ceremonies are all tightly integrated. The physical space itself is the main semantic organizing principle. Visual Immersion provides an **egocentric view**, which allows the observer to view the temple from the inside, as it was meant to be viewed in real life. To our knowledge, ours is the first formal experimental study in the use of immersive virtual reality for virtual heritage.

We chose the game metaphor for the advantages described in section 2.4.6, p38. With the temple, we saw a design opportunity in the information structure of the temple and supporting materials. We were able to structure the learning goals and activities in a way that is inextricable from the topic matter itself. See 3.3, p98, for details. To our knowledge, only Winn (2001) structured an Immersive VR learning experiment as a game in a formal learning experiment with. Our study would be the second.

The most difficult and important goal of our study was to demonstrate how an immersive display could have more utility than a cheaper desktop monitor in a realistic situation. Several previous studies failed to do this (Moreno, 2002b; Byrne, 1996; Rose, 1996; Salzman, 1999) and only one succeeded (Salzman, 1998). Guided by Salzman's experiment, we structured our experiment in terms of the effectiveness of an egocentric view verses an exocentric view, instead of their appropriateness for this particular topic matter. We also thought it important to test for the difference between short-term and long-term retention, and the interaction between students' level of visual skill and the display type.

7.2.2 Gates of Horus

In our study, eighty-five middle-school students (grades 6-8) learned about ancient Egyptian art, religion, and society by playing an educational learning game, *Gates of Horus*. The game is based on a Virtual Egyptian Temple (Jacobson, 2005e, 2004a), which has no real-world analog. Instead, it embodies only the key features of the typical New Kingdom period Egyptian temple in a way that an untrained audience can handle. The temple has four major areas, and each one has a copy of the High Priest, a pedagogical agent.

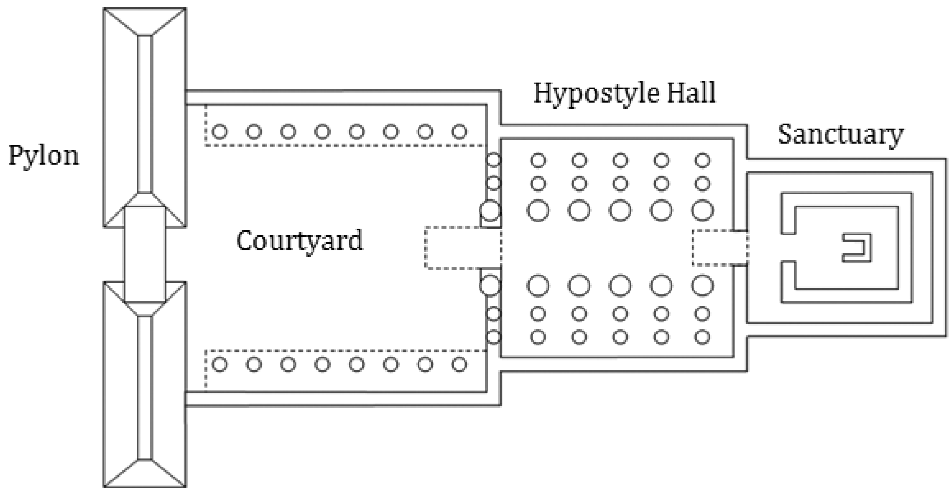

Figure 47. Map of the Virtual Egyptian Temple

Ordinarily, the student navigates the temple in Desktop VR (standard desktop PC) and selects spotlighted features of the Temple. Each time the student selects a feature, the priest will explain what it is and what it does. When the student clicks on the priest, the priest asks the student a question about one of the features the student selected earlier. When the student correctly answers all of the questions for a particular area, the gateway to the next area opens. The goal of the game is to reach the innermost area, the Sanctuary which contains the divine image of the God, Horus. When the student answers all the priest's questions for the Sanctuary, the divine image will speak and bring the blessings of heaven to the land of Egypt.

The software for *Gates of Horus* is based on a commercial game, UT2004 (EpicGames, 2008), and two freeware packages, CaveUT and VRGL (PublicVR, 2008). The freeware enables the game to operate in a variety of immersive displays. In our study, some students saw the temple on a standard desktop monitor while others used the Earth Theater at the Carnegie

Museum of Natural History in Pittsburgh (Figure 48). Otherwise, all students interacted with the game in the same way, one at a time and with a *Gyromouse*. In our study, students navigate through the temple and select items using the *Gyromouse*, a cordless hand-held device used in some VR applications (Duncan, 2006; Herpers, 2005; Olwal, 2002; Patel, 2001; Hafner, 2000; Winograd, 2000).

Figure 48. The Temple in the Earth Theater

7.2.3 Hypotheses

We randomly assigned each student to one of three groups. Each student in the *Theater Group* played *Gates of Horus* using the immersive dome display. Each student in the *Desktop Group* played the game on a standard desktop computer in an area adjoining the main theater. Members of the *Control Group* also played the game on a standard desktop but took the Post Test for basic knowledge *before* playing the learning game. The Control Group's scores on the Post Test take the place of a knowledge pretest for the other two groups. This is necessary, because any question-and-answer pretest given to the Theater and Desktop Groups would reveal too much information about the temple itself.

Our first *a priori* experimental hypothesis is that (**H0**) students will enjoy playing *Gates of Horus* and engage with it fully during the experiment. The second is that (**H1**) students who play *Gates of Horus* will learn something. A positive result in both hypotheses would show that we made an effective learning tool that students will actually use. Both are prerequisite for our core hypotheses, (**H2**). students who play the game with Immersive VR will learn more and retain more than those who play it using Desktop VR, as measured immediately after testing and (**H3**) one or two months later. We specifically designed the experiment to test these hypotheses.

However, we added three other measures to test hypotheses which were not primary, but worth investigating. They are (**H4**) Students who played the game with ImmersiveVR will have better spatial knowledge of the temple than those who used Desktop VR, (**H5**) One or two months after playing *Gates of Horus*, students using ImmersiveVR will report more motivation to learn about Egypt, and (**H6**) students with low visual reasoning ability will benefit more from visual immersion *or* the opposite will be true (Bricken 1990). Regardless, we expect visual reasoning ability to matter.

7.2.4 Protocol

Upon arrival, each student was randomly assigned to an experimental group, Control, Desktop, or Theater. The following sequence describes the important tasks each student performed during the study.

1. All students completed Raven's Progressive Matrices test, which measures current visual reasoning ability (Raven, 1957; Shiply, 1949; Gregory, 1999).
2. Each student in the Desktop Group played Gates of Horus on a desktop computer.
3. Each student in the Theater Group played Gates of Horus in the Earth Theater.
4. Each student in the Control Group took the Post Test before playing Gates of Horus.
5. Post Test questions are standard multiple-choice or short-answer.
6. Gates of Horus records everything during play.
7. Each student in the Desktop or Theater Group took the Post Test.
8. Each student in the Control Group played Gates of Horus on a desktop computer.
9. Each student drew a map of the temple to test his or her knowledge of its layout (Drawn Map Test).
10. Each student placed magnets representing features of the temple onto a provided map of the Temple (Magnet Map Test).
11. Each student produced a video tour of the virtual temple. The student stood in front of a projection of the virtual Egyptian temple, and navigated using the Gyromouse. The student conducted a show-and-tell tour of the virtual temple, which we recorded as a simple documentary video.
12. One or two months later, the student completed the Follow-Up test via the Internet using any computer with a Web browser. This test is also a standard quiz, with multiple-choice and short-answer questions.

7.3 RESULTS

7.3.1 Acceptance

In the **Post Test,** students in the Theater and Desktop Groups answered a first set of questions about their **opinions and feelings** toward the learning game. (Students in the Control Group skipped these, because they had not yet played the game.) Results were strongly positive for all questions. Figure 49, below, illustrates results for the two most important.

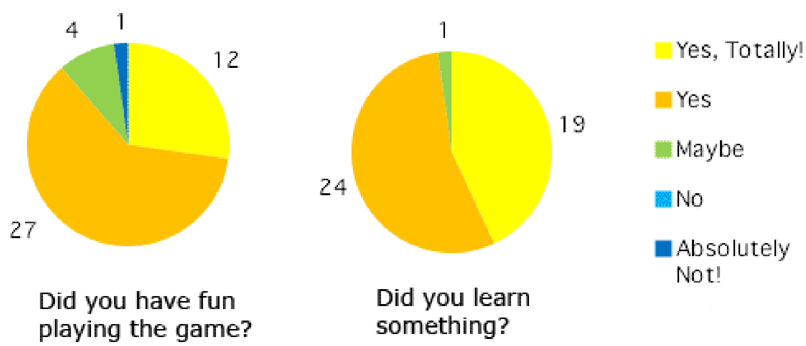

Figure 49. Key Affective Results

In addition to giving these strong self reports, nearly all students played the game from beginning to end, 45 minutes to one hour, without even asking for a break. Together, these factors indicate the students were fully engaged with the learning task (Gates of Horus) during our study, which satisfies **H0**. Because we did not ask the students to compare the game with anything else, this is not a *statistical* proof. It would be interesting to compare Gates of Horus to other learning games and other methods of instruction, but that is beyond the scope of this study.

7.3.2 Effectiveness

The remaining bulk of the **Post Test** quizzed students on their knowledge of the temple. We gave the test to the Control Group *before* they played the game to determine how well students could guess their way using prior knowledge and employing test-taking techniques. All other students played Gates of Horus first, and we compared their Post Test results with those of the Control Group.

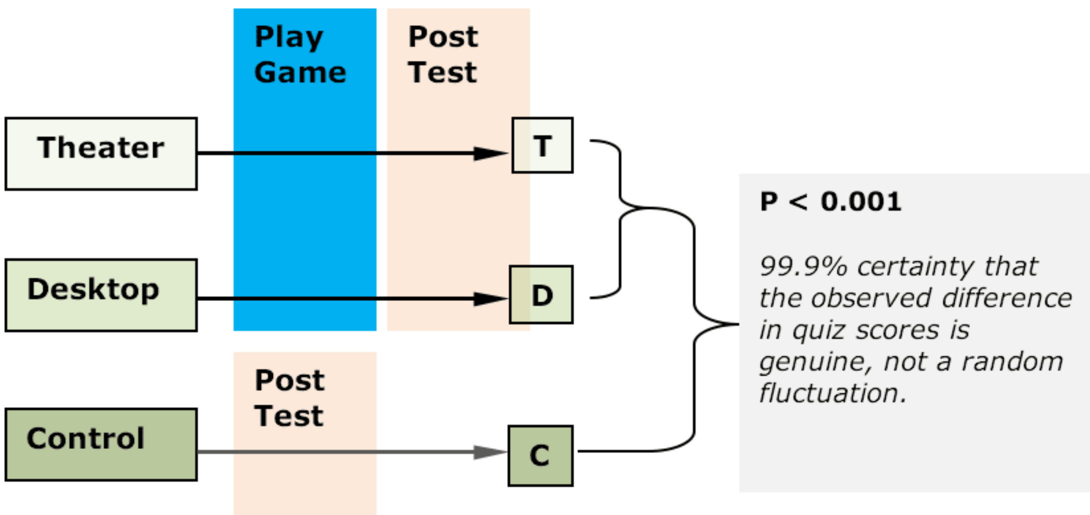

Figure 50. Post Test Sequence and Results.

Figure 50 shows the portion of our testing sequence which pertains to the Post Test. Each student earned a total score awarded by four graders (section 4.4, p144). The Post Test scores produced by students who played the game showed a high probability of being different from the scores of those who did not ($P < 0.001$) when analyzed with the Mann-Whitney statistical test. Furthermore, students who played the game did (statistically) significantly better ($P < 0.05$) in 22 out of the 45 relevant items on the Post Test. This is strong evidence that Gates of Horus is an effective learning tool, which satisfies **H1.** It allows us to meaningfully make further comparisons on the effectiveness of different versions of the game.

Next, we compared Post Test scores for the Theater Group versus Post Test scores for the Desktop Group **(H2)**. The Theater group seemed to do better, but the difference was not statistically significant. We believe that a quiz of this type is not sufficiently sensitive to measure the conceptual learning advantage that we expect from visual immersion.

7.3.3 Immersion

Each student produced a documentary video. First, three human graders evaluated the video for factual knowledge, and their judgments were combined to produce a final score. Our most important result was a significant difference between the *Video Test* results from the Theater Group with results for the other two groups combined. Figure 51 shows the stages of our testing sequence relevant to the Video Test and its main results. We included the Video Test scores from the Control Group, after they had played the game, because we did not think that the order of the Post Test and the game would make a substantial difference in their overall learning.

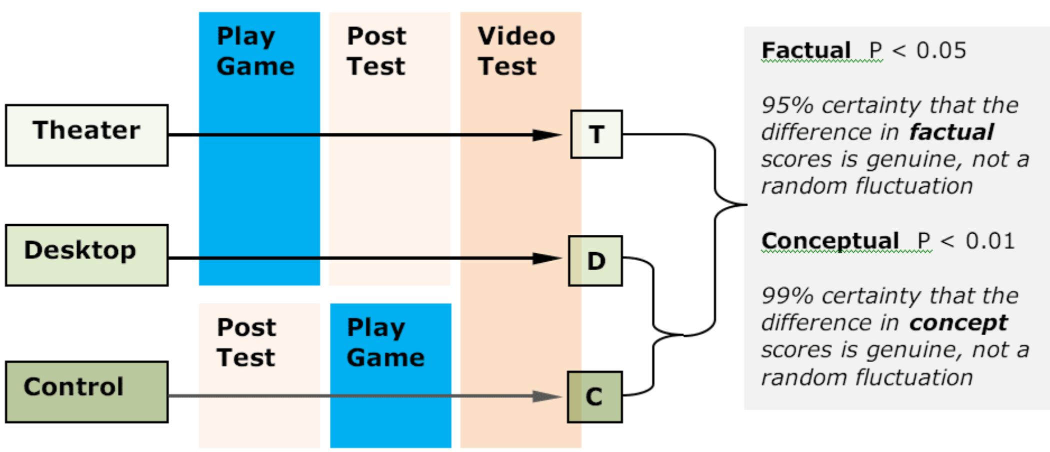

Figure 51. Video Test Sequence and Results.

Statistically, students in the Theater Group did significantly better than students in the other two groups (Mann-Whitney test, $P < 0.05$). Looking at the individual facts that the students were expected to master, the Theater Group did significantly better in 9 out of 45 at the $P < 0.05$ level and we found that 14 out of 45 at the $P < 0.10$ level. Taken together, these results support **H2**.

Next, our Egyptologist, Dr. Lynn Holden, evaluated each student's video for the student's mastery of conceptual knowledge about the temple. Here, the difference was statistically much stronger, and a respectable number of individual measures also showed results significantly different (in 6 of 19, $P < 0.05$). Dr. Holden also assigned an overall impression of the student's mastery of the information, which provided the strongest difference of all, $P < 0.001$. Not only were the main results for the conceptual and factual measures in accord, but they also tended to parallel each other in the individual measures. For every concept where the Theater Group statistically did better than the others, there was at least one fact related to that concept where the Theater Group *also* did better. We are confident that this data satisfies **H2**.

7.3.4 Other Measures

We also compared results for the Follow-Up Test, one or two months later, and comparing responses from students in the Theater Group to all the others. Unfortunately, we saw no statistically significant differences, and there were not enough differences in individual scores to be convincing. The Follow-Up Test was very similar to the Post Test, except that students gave much less informative answers on the short-answer questions. We conclude that if there are any lasting comparative benefits to visual immersion over a standard desktop monitor, a standard quiz such as a Follow-Up Test is not sensitive enough to detect it. This leaves **H3** unproven.

In data from both the Drawn Map Test for **H4** and the Magnet Map Test for **H5**, we saw no significant differences in performance for the Theater Group versus the other groups, not even a trend in one direction or the other. This is at odds with the virtual reality training literature, which has established that Immersive VR is a good way to teach survey and route knowledge of an area. We conclude that the temple was too small and too simple for immersion to produce a genuine difference in students' knowledge which could be detected with a mapping test.

7.3.5 Visual Reasoning Skill

We did see an interesting interaction between students' visual reasoning ability and their overall rating in the conceptual video data. As expected, we found that students with higher RPM scores did significantly better, overall, than students with low RPM scores ($P < 0.05$). More interestingly, students with low RPM scores seemed to benefit more from visual immersion than those with higher RPM scores. Table 35 summarizes our findings.

Table 35. Summary of RPM Effect

	Immersion (Theater Group)	Non-Immersion (Desktop Group & Control Group)	Mann-Whitney test
Low RPM Score	23 students' scores for the Video Test	23 students' scores for the Video Test	P = 0.01
High RPM Score	24 student's scores for the Video Test	24 students' scores for the Video Test	P = 0.09
High RPM vs. Low RPM	All students in both conditions		P = 0.05

As we describe earlier, our Egyptologist evaluated students' video tests for conceptual knowledge. Among other measures, he assigned each one an overall rating for mastery of the material. This is the conceptual score referred to in Figure 51. **Video Test Sequence and Results.**, p253. Using the Mann-Whitney test, we saw that scores for Low-RPM students in the Theater Group were significantly different from the scores of those in the other groups. Differences at the 0.01 level are considered quite strong. There was also a difference for High-RPM students, at the $P = 0.09$ level, which is considerably weaker. Although it is far from conclusive, we consider this evidence to support **H6**.

Our evidence for **H6** implies that students with lower (current) visual reasoning ability benefit more from visual immersion than their more visually skilled classmates. One other study showed an interaction between visual reasoning and learning with virtual reality. In Winn (1997), students with low RPM scores benefited more (than students with high RPM scores) from producing small virtual environments than from experiencing them. Our results and his imply that further research comparing visual reasoning skill/ability and learning results in Immersive VR may be fruitful.

7.4 DISCUSSION

Our study directly demonstrates that a visually immersive display is more effective than a standard computer monitor for a reasonable learning activity in an important topic area (**H2**). The immediate implication is that Immersive Virtual Reality is the optimal interface for learning declarative and conceptual knowledge, if:

1. Some visual artifact effect or process provides the central organizing theme for the material.
2. The immersive interface allows the student to interact with the information in some way not otherwise available or efficient.
3. Almost every way in which the student interacts with the virtual environment is relevant to the learning task.

In our study, a very wide egocentric view of the Virtual Egyptian Temple allowed students to (1) navigate less and think more and (2) see the interior of the temple as a whole, rather than as a collection of features. Generally, if the objective is for the student to understand some artifact system or phenomenon as a whole, have the student work with it as a whole.

In some topics, Immersive VR is the best medium for this. Salzman (1998) produced the only other study to clearly demonstrate an educational advantage of adding Immersive VR's capabilities to what a non-immersive display can provide. Her experiment centered on understanding magnetic fields and science topics which can be presented in a similar way. By achieving a similar result in the area of cultural heritage, we have identified another large intellectual territory where Immersive VR has practical use. We hope that future studies will provide more such working examples to add to the existing literature. For example, someone could conduct an experiment in which students learn to "read" an urban neighborhood to determine its planning needs.

Our data showed that the immersed students learned better than those with only a monitor, but it did not tell us *why*. Perhaps the immersed students (a) had a mechanical advantage to access information encoded in the environment, (b) developed a deeper understanding of the topic materials, because the immersive view presents the information in a more coherent manner (c) experienced a deeper sense of presence in the virtual environment or (d) benefited from some mixture of all three factors. We can safely discount *presence* as a significant factor because self-reported presence was only slightly different between students who used Immersive VR and those who use desktop VR. Furthermore, previous studies which isolated presence in ImmersiveVR as a potential influence on learning showed no effect (Moreno, 2002b; Byrne, 1996; Rose, 1996; Salzman, 1999). Separating the other two factors would not be so simple.

We could certainly restructure the study to isolate the benefits of the Immersive view from the benefits of the physical interaction. That would require a difficult and interesting set of experiments concerned with understanding how the computer and the human work together. Alternatively, we could try a "brute force" approach, where we repeat the experiment with better evaluation tools, a larger sample size, a longer exposure time (perhaps as part of a larger curriculum) and a larger and more meaningful lesson. If successful, this would further verify our results and allow us to make fine-grained comparisons of factors such as the effect of RPM score, gender, or video game experience.

We are particularly interested in our data which *implies* that students with lower visual reasoning skills (as measured by RPM score) benefited more from visual immersion in our experiment (**H6**). It would be fairly straightforward to conduct follow up experiments to see if this is really happening. If the finding is corroborated, perhaps Immersive VR could be used to improve visual reasoning skills among the students who need it most. In the meantime, it could help them keep up with their more skilled peers in certain topics.

We could deepen the study by adding more interaction possibilities for the student. We are particularly interested in adding people to the Virtual Egyptian Temple (or some other space) to provide a social context. These people could be a mixture of automatons, educators, and other students, re-creating aspects of the ancient society. As a teaching tool, adding virtual people should be very effective, and for the archeologist it would be a means of testing hypotheses on what ancient cultures were like.

We are pleased to see that the students liked the game, Gates of Horus, which was our central learning activity for all conditions and experiment (**H0**). The fact that students did learn from the game, regardless of interface, is also important and useful (**H1**). Without these results, data comparing learning with or without immersion (**H3**) would have been meaningless.

Our ultimate goal is to inform the educator on when and how to employ Immersive Virtual Reality. We see it as a new learning tool which the educator and student can combine with existing approaches to solve real problems for real people. With that in mind, we look at these trends:

1. Educational research applications and practice using Desktop VR have exploded in recent years, with the bulk of the growth in shared online communities. Students' comfort with technology and continuing advances in personal electronics and telecommunications make Desktop VR increasingly practical.
2. The sophistication of projection hardware is increasing as its cost is decreasing, which has made visually immersive displays affordable for schools, museums and small institutions. A very important part of this trend is the fact that planetaria and dome displays are increasingly all-digital. Eventually, personal visual immersive displays will become affordable for individuals in schools.
3. The gaming metaphor is becoming increasingly influential in education, as educators exploit the opportunities it affords. Students' high acceptance and intense concentration on current video games is irresistible to educators. It will become more important, as new teachers who played video games as children grow up to be teachers.
4. A wide range of industries and academic disciplines are increasingly adopting computer graphics of all types. Within almost every professional literature, one can find a small but persistent subset concerned with the use of virtual reality to address their topic.
5. Museums, especially science museums, continue to innovate with interactive displays of all types. Unlike the schools, they must make their exhibits and educational programs interesting to continue to attract patrons.

We believe these trends are converging toward a wider use of interactive immersive media, especially in museum education and educational gaming. We believe that our study was well-positioned to explore questions relevant to this near future, and we look forward to building on the results.