Building and Using Educational Virtual Environments 
for Teaching About Animal Behaviors

A Dissertation 
Presented to 
The Academic Faculty

By

Donald Lee Allison, Jr.

In Partial Fulfillment 
Of the Requirements for the Degree 
Doctor of Philosophy in the College of Computing

Georgia Institute of Technology 
December, 2003

Copyright © 2003 by Donald L. Allison, Jr.
Approved by:

Dr. Larry F. Hodges, Advisor

Dr. Chris Shaw

Dr. Mark Guzdial

Dr. Jean Winebr

Dr. Blair MacIntyre

Date Approved: 20 November 2003
This dissertation is dedicated to the memory of my mother. She always believed that I could do it even when I wasn't so sure.
The process of completing a Ph.D. is a long, arduous one. Even though the research is conducted by a single individual, there are many others who fulfill important roles in the process. Thanks are due first to my advisors, whose positive attitude and encouragement were crucial to the completion of this dissertation. Thanks are also due to the virtual environments group members upon whose shoulders I have stood. I'd still be trying to build my virtual environment if I hadn't had the SVE toolkit available. My committee played an important role in the process, getting me to tons down the proposal and quit trying to solve everything at once, while at the same time focusing on the actual accomplishments of this work.

Then there are those who helped in less direct but no less important ways, notably the staff of the College of Computing. My thanks to those in the accounting, student services, main, and GVU offices, who went above and beyond the call of duty many times.

Special thanks are due to Sandy Shew and Rob Melby, for their help with Perl scripts, especially on a PC. Finally I'd like to acknowledge the moral support of my family, all of whom already have Ph.D. degrees. Their faith and encouragement helped me through some rough times.
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedication</td>
</tr>
<tr>
<td>Acknowledgement</td>
</tr>
<tr>
<td>Table of Contents</td>
</tr>
<tr>
<td>List of Tables</td>
</tr>
<tr>
<td>List of Figures</td>
</tr>
<tr>
<td>Abbreviations</td>
</tr>
<tr>
<td>Summary</td>
</tr>
<tr>
<td>Chapter 1. Introduction</td>
</tr>
<tr>
<td>1.1 Introduction</td>
</tr>
<tr>
<td>1.2 VR Successes</td>
</tr>
<tr>
<td>1.3 VR and Education</td>
</tr>
<tr>
<td>1.4 Cognitive Science Support for Educational VR</td>
</tr>
<tr>
<td>1.4 Thesis of this Dissertation</td>
</tr>
<tr>
<td>1.5 Organization of this Dissertation</td>
</tr>
<tr>
<td>Chapter 2. Related Work</td>
</tr>
<tr>
<td>2.1 Concept Acquisition Using VR</td>
</tr>
<tr>
<td>2.1.1 Breisford</td>
</tr>
<tr>
<td>2.1.2 ScienceSpace</td>
</tr>
<tr>
<td>2.1.3 VRRV</td>
</tr>
</tbody>
</table>
Chapter 3. Building the Virtual Gorilla Environment

3.1 Gorilla Identification
3.2 Gorilla Vocalizations
3.3 Zoo Habitat Design
3.4 Gorilla Dominance Hierarchy
3.5 Gorilla Social Interactions
  3.5.1 Gorilla Motions
  3.5.2 Gorilla Behavior Controller

Chapter 4. Testing the Virtual Gorilla Environment

4.1 Qualitative Testing
4.2 Prototype Revisions
4.3 Quantitative Evaluation
LIST OF TABLES

Table 1-1. Number of VR Papers Published by Year
Table 2-1. Constants for Computing Gorilla Limb Lengths
Table 2-2. Typical Limb Lengths for Adult Gorillas
Table 2-3. Gorilla Ratios of Limb Lengths (Based on 21 Adult Males)
Table 2-4. Male Gorilla Limb Lengths from Limb Proportions
Table 2-5. Working Values for Male Gorilla Gorilla Limb Lengths
Table 2-6. Gorilla Limb Circumferences
Table 2-7. Body Measurements for Young Gorillas
Table 3-1. Statistical Analysis of the Two Tests
Table 3-2. Mean and Standard Deviation of (Posttest B - Pretest A)
Table 3-3. Confidence Interval and P Value for Hypothesis $\mu_{\text{Post B}} = \mu_{\text{Pre}}$
with Pretest A and Posttest B
Table 3-4. Test Task Descriptions
Table 3-5. Mean and Standard Deviation of (Posttest A - Pretest B)
Table 3-6. Confidence Interval and P Value for Hypothesis $\mu_{\text{Post A}} = \mu_{\text{Pre}}$
with Pretest B and Posttest A
Table 3-7. Mean and Standard Deviation of Posttest A Results
Table 3-8. Confidence Interval and P Value for Hypothesis $\mu_{\text{Post B}} = \mu_{\text{Pre}}$
with Posttest A
Table 4-1. Mean and Standard Deviation of Posttest B Results
Table 4-2. Confidence Interval and P Value for Hypothesis $\mu_{\text{Post B}} = \mu_{\text{Pre}}$
with Posttest B
Table 4-11. Mean and Standard Deviation of Test A Results

Table 4-12. Confidence Interval and P Value for Hypothesis $\mu_{\text{actual}} = \mu_{\text{predicted}}$

Table 4-13. Mean and Standard Deviation of Test B Results

Table 4-14. Confidence Interval and P Value for Hypothesis $\mu_{\text{actual}} = \mu_{\text{predicted} \text{NR}}$
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1</td>
<td>Modeled vs Actual Silverback</td>
<td>40</td>
</tr>
<tr>
<td>3-2</td>
<td>Modeled vs Actual Female</td>
<td>40</td>
</tr>
<tr>
<td>3-3</td>
<td>Modeled vs Actual Habitat</td>
<td>44</td>
</tr>
<tr>
<td>3-4</td>
<td>Behavior Controller Overview</td>
<td>51</td>
</tr>
<tr>
<td>4-1</td>
<td>Prototype System Test Setup at Zoo Atlanta</td>
<td>56</td>
</tr>
<tr>
<td>4-2</td>
<td>Pretest Results Distributions</td>
<td>62</td>
</tr>
</tbody>
</table>
ABBREVIATIONS

AAPT: American Association of Physics Teachers
ACM: Association for Computing Machinery
AERA: American Educational Research Association
BOOM: binocular omni-oriented monitor
C ACM: Communications of the ACM
CAVE: CAVE automatic virtual environment
CG&A: IEEE Computer Graphics and Applications
EVL, Electronic Visualization Laboratory
HITL, Human Interface Technology Lab
HMD: head mounted display
ICLS: International Conference of the Learning Sciences
IEEE: Institute for Electrical and Electronic Engineers
IVE: immersive virtual environment
VE: virtual environment
VR: virtual reality
SUMMARY

In order to advance the technology and drive further development, applications need to be found where virtual reality is beneficial, or at least not detrimental. Flight simulators and phobia treatment are two such applications, but they are niche applications with very specific audiences. Education has been proposed as an application area where VR might be broadly applicable and have far reaching effects, but studies of this area are still in their early stages. This dissertation examines the learning potential of the virtual gorilla environment.

First, other educational virtual environments are examined to determine what lessons can be learned from them. Since an important aspect of the virtual gorilla environment is virtual gorillas that interact as real ones do, other animal simulations were evaluated for useful ideas for building virtual gorillas.

A simulation of the gorilla environment of habitat 3 at Zoo Atlanta was constructed, emphasizing educational goals designed jointly with Zoo Atlanta. Two groups of students were given pre- and post-tests on gorilla identification, behavior, and social interactions, with one group exploring the virtual habitat between tests. A statistical analysis showed that student performance improved on some topics after exposure to the virtual gorilla environment, but not others.
CHAPTER 1
INTRODUCTION

Introduction,

Virtual Reality is a technology in search of a defining application. With any new technology, there are people willing to experiment with it as soon as it is announced. These people are willing to build their own equipment and to endure the high price and poor performance of the technology as implemented initially. It is not until the market demand is there, though, that it becomes economically feasible to invest large sums of money in research and development to improve the technology, and to enable it to be mass produced in a cost effective manner. This doesn’t happen until someone devises a defining application, one that is so compelling that many people are willing to put up with the initially poor implementation of the technology in order to be able to use that application.

A recent example of this is 3D computer graphics technology. Once the domain of million dollar Silicon Graphics computers, some pioneers had the dream of bringing that capability to personal computers. Initial PC graphics cards were either monochrome, or boasted such resolutions as 320x240 with 4 colors, or 640x480 with 16 colors, and had only limited 2D functionality. 3D graphics cards eventually became available for the PC architecture, but costing an order of magnitude or more than the cost of the rest of the computer, and with poor performance.

With the introduction of Doom and other 3D first person shooters, there was a large group of new potential users of 3D graphics. Initial games relied on software rendering to draw pictures using 2D graphics cards and made many simplifying
assumptions in order to be able to draw anything that looked mildly three dimensional. These games drew a large enough audience, though, that it started to become economically feasible for chip manufacturers to develop hardware support for some limited 3D functionality. Game developers were able to take advantage of this extra functionality in their new games, each of which had to surpass previously introduced ones in order to capture market share.

This cycle of game developers working with and pushing the 3D graphics card manufacturers, who added more functionality and challenged game developers to take advantage of it resulted in high performance PC graphics at a price comparable to the cost of the rest of the computer instead of costing ten or more times that much. Driven by market demand, for a few hundred dollars it is possible to buy a graphics card that will display several tens of millions of lit, shaded triangles per second. 3D graphics capabilities are now available for anyone to experiment with, at the cost of a few hundred dollars, and have moved into the mainstream market. Along with becoming more affordable, absolute performance has improved as well so that today's PC is a much more capable performer than the SGI Reality Engine of a few years ago.

For the case of 3D graphics functionality, first person shooter computer games were the driving application that brought the technology to the masses. Virtual reality, unfortunately, has not found such a defining application yet. Immersive virtual environments are currently experienced using narrow field of view, low resolution head mounted displays (HMD's), or stereo glasses and one or more walls of large, back projected video screens which end up dimly lit and with low effective resolution. These displays are expensive, cumbersome, and have many ways of reminding the user that he is not actually in the virtual world. Similarly, tracker technology, needed to determine
gaze direction for HMD's and body motion for general interaction is intrusive, cumbersome, expensive, has low repetition rates, and large lag.

Although there has been some research on providing stimuli for other senses such as touch or smell, this work is at an even earlier stage than the work on optical stimulation. The problem is that current VR applications, such as flight simulators or entertainment emporiums such as Dave and Busters\(^3\), provide only a limited market for VR technology. Without a larger market demand, manufacturers are reluctant to invest the millions of dollars necessary for the technology and manufacturing breakthroughs needed to reduce the component price while improving performance and addressing the current technological drawbacks, since they see a negative return on investment.

A few companies have tried to be market leaders, attempting to catch the leading edge of what they perceived as the untapped market for VR technology. They have introduced VR components at prices ranging from several hundred dollars (the i-glasses tracked, head-mounted display\(^1\)) to a few thousand dollars (the VFX3D tracked HMD\(^2\) or the Sony Glasstron HMD\(^3\)) to several tens of thousands of dollars (the Virtuality HMD and hand tracker gaming system\(^4\)). These companies have either uniformly abandoned the market, or else gone out of business or bankrupt trying to drive it. No one has yet developed an application so compelling that users are willing to pay the high prices and endure the low quality of the technology currently available in order to have personal VR. Until the market demand is there to drive the research and development, virtual environments such as those described in *Virtual Destruction*\(^5\) or *Society of the Mind*\(^6\) will remain in the realm of science fiction.

From the above, it should be apparent that an important question to explore is to determine for which applications VR technology is a useful tool, and for which it contributes nothing of value. Once it is known where VR technology is potentially useful,
those applications can be investigated and developed further to determine their potential ubiquity.

VR Successes.

Fred Brooks has provided a summary of the state of VR in his 1999 CG&A article7. As he noted, the best VR application available today is the flight simulator, although the flight simulator community prefers not to associate itself with the field of virtual reality. Today's flight simulators are so effective that pilots can qualify on a new type of aircraft strictly by flying the simulator, without ever having set foot in a real cockpit. By using actual dials, knobs, gauges and displays to provide the near-field haptics of the flight controls, using a motion platform to provide proprioceptive feedback, and using computer graphics to generate the views out the aircraft windows, an effective immersive environment has been created. Although each simulator costs millions of dollars and is good only for a single type of aircraft, it is more cost effective than taking an airplane out of service for training pilots. In addition, it is possible for pilots in the simulator to experiment with situations too dangerous to try in a real airplane, and by making mistakes and "crashing", to develop appropriate reflexive responses to these situations for the rare occasion in which they might encounter one while piloting an actual airplane.

Another area described by Dr. Brooks in which VR is having a commercial success is psychotherapy, and in particular, phobia treatment. VR systems have been used successfully to treat such phobias as fear of heights, fear of flying, fear of spiders, post-traumatic stress disorder, and fear of public speaking, and VR systems for such treatments are beginning to move out of the lab and into therapists’ offices.
Both flight training and phobia treatment, however, are specialized applications for which the commercial need is limited to a small number of systems. Neither one will be the defining application for which everyone will want a personal VR system in their home. In fact, one of the problems with commercialization of VR for the treatment of phobias is to find an HMD manufacturer that will still be in business and continue to support its products for more than a couple of years. What is needed is a more ubiquitous application that provides a market for millions of systems instead of tens or hundreds of them.

VR in Education

An application area which has been suggested as one to which VR might make a great contribution is that of education. There is continued concern for the quality of education provided by the American educational system. For the last decade, people have been speculating that VR might either revolutionize or ruin education. This speculation engendered a lot of discussion, which reached a peak in mid-1990’s (see Table 1-1.) and continues today with the token mention in the issue of Computer Graphics focusing on VR. Advocates of VR portrayed the future as allowing each student to learn in the most efficient manner possible, using a VR system specifically tailored to his style of learning. Detractors pointed out that no computer had been able to even approximate the adaptability of humans, and therefore teachers were still the best educational technology yet invented. Just as it was becoming apparent that the initial hype was greatly overdone, and that it would be a very long time before any of the optimistically predicted VR educational systems were in widespread use, the Internet and World Wide Web emerged on the scene and became the new technology that was going to solve all the problems in education. Attention shifted to this new panacea, and
VR researchers were able to explore educational uses of VR with less danger of succumbing to the hype and public hysteria inflamed by the press.

With the increased interest in and research about the web and its educational potential has come a corresponding decrease in the number of people studying educational possibilities of VR. Veronica Pentelidis at East Carolina University has maintained a list of publications related to VR and education starting in 1991\textsuperscript{13}. This list has been available through the internet, and the latest version of the list now states that it has last been updated in April, 2001. This bibliography is considered to be one of the most comprehensive lists by many VR researchers\textsuperscript{14}. The bibliography is broken down by various types of educational uses of VR. Analyzing the publication dates of the references in the list yields the data of table 1.

**Table 1.1: Number of VR Papers Published by Year**

<table>
<thead>
<tr>
<th>Year</th>
<th>General</th>
<th>Collaborative</th>
<th>Disabilities</th>
<th>Medical</th>
<th>Museum</th>
<th>Training</th>
<th>Military</th>
<th>Textual</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>1993</td>
<td>10</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>14</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>25</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>20</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>14</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>18</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>21</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
From this table it is apparent that the initial euphoria over using VR as an
educational technology peaked around 1994, and the interest has consistently waned
since then. However, the waning interest is not because of empirical evidence that VR
is not useful in education, but because answering the question is going to be a long term
process in which the answer is not an unequivocal "yes" or "no," but a "maybe"
conditioned by certain conditions and only restricted to specific subject areas.

Representative of the two opposing points of view as to the efficacy of VR as an
educational technology was the debate between Fred Brooks, Kenan Professor of
Computer Science at UNC Chapel Hill, an early computer science and VR pioneer, and
Chris Dede, then of George Mason University and currently chair of Learning and
Teaching at the Harvard Graduate School of Education, an educational technologist.
Brooks contended that VR had nothing to contribute to education that couldn't already
be done cheaper and more efficiently with other techniques. Specifically, he argued "...I
see little performance/price advantage for VR over other educational technologies for
college education, much less secondary; it is rank foolishness to talk about it for
elementary education." In an interview with Howard Rheingold, he also expressed
concern about virtual environments actually misleading and misinforming people:

The potentials for misleading are very great in that kind of
instance...fractal mountains are a good clear visual image of an important
distinction between realism and truthfulness. The danger of more and
more realism is that if you don't have corresponding truthfulness, you
Teach people things that are not so. In business scenarios, or war games,
to the extent that your model of the business world or the war world is not
real, you can make the mistake of teaching people very effectively how to
apply tactics and strategy that won't work in the real world.

Chris Dede, on the other hand, contended that VR could be very useful by
providing early educational experiences that children otherwise wouldn't have. Children
could experience and interact with a world in which the quantum constant was large
enough that its effects on object interactions would be quite noticeable. Then when
these same children were introduced more rigorously to quantum physics in college, they would already have an intuitive foundation upon which to build their knowledge of quantum interactions.

Cognitive Science Support for Educational VR

Cognitive science does provide theoretical support for the application of VR technology to education, so proponents are not simply engaged in wishful thinking. Brer summarized the cognitive science approach to learning in his book, *Schools for Thought*. Learning is seen as a process of moving from novice to expert problem solver. One key component is the combining of units of knowledge together to form larger chunks, which are represented with a single symbol and are treated as a single entity. Since the mind has a limited capacity to store and manipulate symbols, the more powerful one's symbols are, the better that person can reason.

A student's knowledge is encoded in if-then production rules, which are used when the student reasons to solve a problem. The better a student's production rules, the more expert a problem solver he is, and the more he is said to have learned. Learning is the process of acquiring more accurate and more general production rules. This occurs when the student is exposed to a problem for which his current production rules do not provide the right answer. By refining his rules, the student is able to handle a larger class of problems. To facilitate learning, then, a student should be exposed to a wide variety of problems of a particular type.

Sometimes, though, a student is not able to solve a particular problem because the production rules encoded in his knowledge base ignore a key data element. For instance, shown a seesaw and asked which side would lower to the ground, the student might only concentrate on the weights on both sides, and ignore the distance of the
people from the pivot. Using just weights, it is impossible to correctly solve seesaw problems where the lighter weight person is far enough from the fulcrum that the resulting torque is larger than that produced by the heavier person. If the student's model just includes weights and not distances from the fulcrum, then there will be some situations in which the student's model predicts one outcome but actual experience shows the outcome to be just the opposite of what was expected. At this point, the student might deduce that other factors were at work as well, and arrive at the conclusion that distance from pivot was also important. However, sometimes the student does not make that connection, and for learning to occur, the teacher needs to focus the student's attention on distance from pivot until the student realizes that that is a key datum.

A virtual environment would appear to be ideally suited to provide exactly this kind of instruction. In a virtual environment it is possible to experience events that are too dangerous to experience in real life. Students can repeat an event over and over as necessary until they have observed all the details of interest. Virtual environments also make scaling possible, so that students can experience events from the past, or events that normally occur too rapidly or too slowly to interact with, such as explosions or the chemical transformation of rusting (when time is scaled) or events that are too physically distant, or too large or too small to interact normally with, such as interacting with the nucleus of an atom, or a galaxy. Once a virtual environment is created, it takes almost no time to set it up and revisit it, unlike many real experiments, where preparation time is longer than the actual educational part. VR could thus provide a host of different problem situations for the student to experience and from which they could learn, including situations that it would be impossible for students to experience in any other way.
In addition, as with any computer simulation, a virtual environment could detect when the student is having problems focusing on the pertinent conditions that form the preconditions of the rules being taught, and could direct the student's attention to those details he is neglecting. This could be done with subtle cues initially, that would become more direct as time passed if the student didn't pick up on the hints.

Virtual reality, therefore, has the possibility to be a useful and powerful educational technology, enabling educational opportunities otherwise difficult to provide. If that turns out to be the case, the educational market is large enough that it could drive innovation in VR technology so that the environments become even more engaging and realistic. Once it became clear that the answer to the question of efficacy was not always a clear cut yes or no, the media hype about VR and education died down. Now researchers are beginning to determine when and where VR is useful as an educational technology. This dissertation examines one facet of the question, studying the use of VR technology to teach about animal behaviors and social interactions.

Thesis of this Dissertation

When and where VR might be useful in knowledge acquisition is still an open question. This dissertation examines one particular application of a virtual environment for education. The problem studied in this dissertation was that of using virtual reality to teach students about animal behaviors and social interactions. The thesis that it investigated was that virtual reality could teach students concepts about gorillas that were hard to convey to students in the traditional manner used by zoos. Zoos place animals on exhibit and use signs to provide information about the animals to the public, and to guide the public's observations. However, there is a body of information that is difficult to convey to visitors in this fashion, due to the distance and intervening materials.
between visitors and the animals, and to the fact that a lot of the more interesting and
educational animal interactions occur off exhibit by choice.

This dissertation will describe a virtual gorilla environment that was constructed
and used to try to teach students about gorilla social behaviors and interactions. It
investigated the conjectures that virtual reality could be used to teach students about the
dominance hierarchy of gorillas; that it could be used to inform students about principles
of gorilla habitat design, that it could teach students to recognize various gorilla
interactions, including vocalizations and postures, used to establish the dominance
hierarchy; and that it could be used to instruct students in the proper identification of
various gorilla types.

Specific contributions of this research include a quantitative evaluation of the
effectiveness of a virtual environment in teaching the four basic areas described above.
Also described in this dissertation as a technical contribution is an architecture for
constructing behavior controllers for gorillas, which could be generalizeable to controlling
the social interactions of other species that have dominance-submission interactions.

Organization of this Dissertation

This chapter has described the problem addressed by this dissertation in the
context in which it arises, and has spelled out the thesis of this dissertation, as well as
providing an overview of its contributions. The next chapter will examine the relevant
literature, summarizing other educational VR systems and placing this work in context. It
will also briefly survey current research into building behavior controllers for computer
controlled animals.

Chapter 3 will describe the virtual gorilla system emphasizing the aspects of its
construction that were targeted specifically at the education goals described in the
previous section. Chapter 4 will provide the specifics of a study undertaken using the virtual gorilla environment to determine how well, if at all, it did teaching students the five areas of its focus: gorilla identification, gorilla vocalization identification, gorilla habitat design, proper gorilla social behaviors, and the gorilla dominance hierarchy. Finally, chapter 5 will examine the results of this study, examine what conclusions might be drawn, and then speculate on the future of virtual reality in educational applications.
CHAPTER 2

RELATED WORK

Flight simulators were just the first application to use a virtual environment for task training. Other applications of VR to task training include training the officer of the deck on a submarine, training shipboard firefighters, teaching palpation of subsurface tumors, training students to use a tree harvester, and situational training of nuclear non-proliferation inspectors and escorts. While interesting applications of VR to education, the focus of the work presented below is the use of virtual environments to teach concept formation and acquisition, or more generally, knowledge acquisition. Therefore task training systems will not be considered further in this work.

By the same token, some people use a very broad definition of virtual reality to include QuickTime VR, web browsing, or even book reading. However, this dissertation is specifically concerned with the effectiveness of immersive virtual environments on concept acquisition (specifically concepts about gorilla behavior), where immersive is defined to include a tracked viewpoint that changes with the user head position. Therefore only systems based on HMDs, CAVES, ImmersaDesks, BOOMs, or their kin, will be considered here.

This chapter reviews other immersive educational VR systems in the sense defined above. It then provides an overview of work on building reasonable virtual facsimiles of real animals.

Concept Acquisition Using VR

Brelsford One of the earliest studies showing the feasibility of using VR for concept acquisition was reported by Brelsford23 in 1993. His research showed
that it was possible to acquire simple physics concepts through interaction with a virtual environment as effectively as these concepts could be acquired through a lecture, and that these concepts were retained as effectively when mastered in a virtual environment as when learned in lecture.

He constructed a simple physics environment, consisting of a pendulum of controllable length and three balls of uniform size but variable mass. There were many variable parameters including air resistance, gravity, mass, friction, pendulum period, and so on. Students were either given an hour lecture, or an hour for a lab in the virtual environment (VE) where they had to solve two problems. Students using the virtual environment were given an initial 20 minute orientation lecture on using the VE, and a 10 minute period in which to study the two problems to be solved, before being immersed in the environment. Both groups were given a pretest, and then four weeks after the experiment were given an unannounced posttest to measure long-term retention of the basic concepts. The experiment was performed twice, once on a group of junior high students and once on a group of college students. In both cases the group using the VE had higher retention than the group receiving the traditional lecture. Thus, exploring a virtual environment given specific problems to solve was as effective at teaching students about the physics of pendula as a traditional lecture.

*ScienceSpace* Another series of studies of the utility of a virtual environment for mastering concepts from physics has been conducted by Bowen Lotin and Chris Dede, among others. Project ScienceSpace has developed over the years into three different virtual worlds. Features common to all three worlds include the use of a head-mounted display for immersion and trackers on the head and both hands for tracking gaze and for interaction with and control of the
environment. Menu control of the environment was provided by menus attached to the non-dominant hand while the dominant hand was used to select menu items by pointing and using a switch. All environments provided a variety of sounds to indicate the types of events occurring in the environment, and a game vest was used to provide elementary haptic feedback.

The earliest environment, NewtonWorld, was intended to teach students Newtonian mechanics, specifically Newton’s three laws of motion and the laws of conservation. The system consisted of two balls constrained to move along a straight line. The balls could collide with each other, and with walls at the ends of the motion track. Initially there was no friction or gravity, although these could be turned on. The user was free to choose from several viewpoints, including the viewpoint of one of the balls. When using the environment, a student would be in it for an hour or more, completing a series of lessons consisting of predicting the outcome of a certain set of conditions, setting up those conditions and observing the actual outcome, and then evaluating their initial predictions in light of the actual outcome.

Being the oldest environment, NewtonWorld has received the most formative evaluation. There were several lessons learned through the progression of this environment that are more generally applicable to all educational virtual environments. One surprising lesson learned was that improving the user interface might actually have a negative impact on learning. Students readily assumed that the size of the balls was intended to convey their relative masses, which would ordinarily be a natural mapping. However, in this case it also reinforced one of the misconceptions that the system was intended to overcome, namely that size and mass correlate directly.
This environment was tested on high school students who had had a year of physics. Even after their physics course, many of the students had varying misconceptions about Newton’s laws and conservation. They enjoyed their exposure to the virtual environment, they thought that being able to examine the environment from many points of view was useful, but there was no statistically significant improvement in their understanding of Newtonian mechanics after their exposure to the virtual environment.

The second environment, MaxwellWorld, was designed to teach students concepts of electrostatics, particularly electric field and electric potential, and to reinforce that these are three dimensional phenomena. The world consisted of a one meter cube in which the student could place positive and negative charges. The student could then explore the resulting electric field by the use of test charges, electric field lines, equipotential surfaces, or flux lines through surfaces.

The first evaluation of MaxwellWorld was performed using high school students who had completed a year of high school physics. Students interacted with the environment in three sessions, each lasting approximately two hours. Pre- and post-testing showed that these students developed a deeper understanding of the distribution of forces in an electric field. Students rated the experience stimulating, and fairly easy to use. Students reported problems when navigating, using menus, and deleting source charges. In addition, some students experienced simulator sickness symptoms ranging from eye strain through stomach discomfort and nausea.

The results of this first evaluation were reincorporated into MaxwellWorld through a series of changes, and a second evaluation was performed, comparing MaxwellWorld to a more traditional two dimensional learning environment, EM Field™. The instructional activities were restricted to those that could be
performed in both MaxwellWorld and EM Field in one part of the study, to compare the differences afforded by working in 3D versus 2D. In a second comparison, both programs were compared using the feature sets of each with no restrictions.

The results for the study comparing 3D and 2D learning environments showed that both significantly improved student conceptual understanding. Students using MaxwellWorld were able to better define the concepts being taught than those using EM Field, although these differences did not persist, as shown by a five month retention test. MaxwellWorld students were better at sketching forces, but EM Field students were better at sketching potential.

Students reported feeling more motivated when using MaxwellWorld, but did not report finding its representations more meaningful. Students using MaxwellWorld reported significantly more simulator sickness and had significantly more trouble using MaxwellWorld. It is worth noting that neither motivation nor interaction experience significantly predicted learning outcomes.

The results of the second phase showed that students demonstrated significantly better conceptual understanding after the lesson, and students who had problems with the concepts being discussed found multirepresentational cues valuable. Motivation, simulator sickness, and usability ratings were similar between the restricted MaxwellWorld from the first phase and the unrestricted MaxwellWorld of the second phase.

PaulingWorld is the newest of the three virtual worlds and is in development. It allows multiple users to interact with a molecule under investigation, changing the orientation of the molecule as well as the representation (currently ball and stick form, van der Waals’ spheres, coded sticks, backbone, or icon replacing repetitive structures). PaulingWorld is
intended to teach students concepts underlying the quantum mechanics of chemical bonds, including probability density and wave functions, ionic versus covalent bonding, and bond angles and bond length.

Overall evaluation of ScienceSpace has led the authors to several general conclusions as to the utility of using a virtual environment to teach concepts. VR features do affect learning, and the quality of the interaction and learning experiences, although VR features can be detrimental to conceptual learning as well as supportive of it. In addition, useability can sometimes support learning and at other times interfere with it. Some factors of a virtual environment are mutually exclusive, forcing the system designer to make design tradeoffs. One such tradeoff encountered by the ScienceSpace authors was that between designing for interaction and designing for learning.

This work, then, serves as a cautionary tale countering the uniform success reported by Breiford described earlier. ScienceSpace has shown that one can't simply build any virtual environment, dump students into it, and expect significant learning to occur. The lessons to be drawn from this work are threefold: first, the content and the audience need to be carefully matched to each other to maximize the possibility of learning occurring.

Second, students find immersive virtual environments to be highly motivating and are willing to interact with them for long periods of time. With a properly matched environment, learning will occur simply due to the prolonged exposure. On the other hand, though, because it involves a large quantity of expensive technology, students expect a virtual environment to be rich and captivating. One recurring theme in the feedback from the physicists at the AAPT meeting where they were given a chance to try NewtonWorld was concern
over the limitations of the environment. They thought it should provide more activities, and that the students should have more control over the environment.

Finally, VR sickness is a real problem with prolonged exposure to a tracked HMD. To minimize the effects of this, lessons should be broken down into units that require use of the HMD in short intervals, instead of one long session, since the resulting headaches and nausea detract from any learning that might occur.

**VR and VR.** The Virtual Reality Roving Vehicle project was a multifaceted project conducted by the Human Interface Technology Lab (HITL) at the University of Washington. One facet involved exposing as many students as possible to preconstructed virtual reality learning environments. A complete virtual environment setup, including head mounted display, head and hand trackers, an interaction wand, and the computers needed to drive the system, were installed into the back of a van, which then set out to visit schools. Upon arrival, students would be given a presentation about VR, and then allowed to explore one of several commercially available virtual environments.

Of more interest pedagogically was a second facet in which students constructed virtual environments representing a topic they were studying. Once they had a basic understanding of their subject, the students planned how to represent the basic concepts in their virtual world. They then modeled the objects needed in their virtual world using Macromedia 3D from Macromedia, and specified how objects interacted together so that the HIT lab programmers could write the necessary code to construct their world. Finally, they experienced their world using the portable VR system.

After the participation of 365 middle and high school students, the VRRV project arrived at several conclusions. Students who participated in building a
virtual world actually did learn the material they were supposed to be modeling, and that occurred regardless of the students' general abilities. Researching, designing, and constructing their virtual world models was the most important factor contributing to student learning.

The VRRV project showed that virtual reality had the potential to reach and teach students for whom the traditional lecture/discussion/homework approach was less successful, and to raise them to a level of understanding comparable to that of their peers for whom traditional education methods are more effective.

NICE: The Narrative-based, Immersive, Constructionist/Collaborative Environments project\textsuperscript{27,31,34} of the Electronic Visualization Laboratory at the University of Illinois at Chicago allowed students to interact with each other as they explored an island and tended a garden on the island. Students could collaboratively plant seeds, drag clouds over to water plants, drag them away to allow the plants to receive sunshine, remove weeds, and thin plants to prevent overcrowding. The plants would provide feedback about its needs to the students through such mechanisms as raising umbrellas when they had received enough water and putting on sunglasses when they had had enough sun. In addition, various computer agents acted as guides to provide suggestions of what to do to the students. Finally, teachers could assume the persona of a guide to provide suggestions to students to further enhance their learning.

The world was persistent and evolving, so that students could revisit it weeks later and see the long term results of their decisions. Students interacted with the world in small groups using a CAVE, with one student being in control of the interactions. Other students joined in from an ImmersaDesk, or via the web.
The system was tested using 52 children six to ten years in age. Students had little problem learning to use the system or mastering the interface. They did find the technology limiting in their ability to orient themselves, which hampered usability and learning. Students were highly motivated and engaged by the environment. Learning correlated with this engagement, so that those students more actively engaged by the environment were the ones who more completely mastered the underlying model of the garden. Students who were in control of the interactions were more engaged than the others, and did correspondingly better on post-tests to measure learning. Pre-testing showed that 12 percent of the students had a good understanding of the concepts in NICE, while post-testing showed an increase to 35 percent of the students. However, of those 35 percent, three-fourths of them had been leaders of their groups, and thus in control of the interactions.

The NICE project showed that the level of a student's engagement is an important factor in determining the amount of learning accomplished via a virtual environment. One problem with the NICE environment was that it was too simplistic. There were only a few concepts contained within it for students to learn, and these concepts were not novel. Students had been exposed to growing plants all their lives and held (possibly incorrect) ideas as to what was involved in raising a garden. A good follow-on project would be one with more content, and content that would be novel for the students, to see if the same results occur again.

Round Earth. The Round Earth project builds upon the work done on the NICE project by the University of Illinois at Chicago. While the cooperative learning in the NICE project was unstructured and undirected, the Round Earth project was more formally defined to meet four criteria: the learning goal needed to be an
important one, one that the educational community agreed was an important part of any educated adult’s pool of knowledge; the goal needed to be difficult, to be one that students ordinarily have problems mastering using ordinary pedagogy; the goal needed to be one that an immersive virtual environment would seem to enhance; and the virtual environment should be based on contemporary educational research. The learning goal chosen to meet these objectives was that of teaching students that celestial bodies such as the earth are approximately spherically shaped, instead of being flat as one’s senses seem to aver.

Two virtual environments were constructed, one a small asteroid and the other based on the earth. In both environments, the object was to pick up ten items scattered either on the surface of the asteroid or in orbit around the earth. Two students would interact in the virtual world together, one playing the role of an astronaut, actually exploring the surface of the asteroid or space around the earth, and the other acting as mission control with a “satellite” view of the world and the ten objects to be found. In the pilot studies, the astronaut would stand in a CAVE while the mission control student would stand in front of an ImmersaDesk. For the formal study, which took place in the school instead of the research lab, the astronaut stood in front of an ImmersaDesk while mission control sat in front of a stereo monitor.

Results from this project were mixed. The first pilot study was used to iron out problems in the system and resulted in changes that improved usability. In the second pilot study, students participated both as mission control and as astronaut, filling each role twice for ten minutes at a time. Students found the system almost too engaging, treating the objective of collecting items as a big video game to the exclusion of focusing on the implicit educational objective. A
the end of the VR activities, students participated in a bridging activity designed to get them to generalize their experiences. Unfortunately, students persisted in maintaining their incorrect mental models, and those who had stated a belief in a spherical earth before the experiment continued to have errors in their mental model.

The limited success of the second test led to a redesign of the educational component of the system. In the third pilot, guides demonstrated the learning objectives to the students and students spent only half as much time immersed in the virtual world. The bridging activity after the immersion used an actual earth globe instead of continuing to use the ImmersaDesk. This third test was again a qualified success. Some students did not improve their mental models, some only temporarily acquired a more accurate model, only to revert to their original model a day or two later, and even those who exhibited persistent learning rarely exhibited it for all of the target knowledge areas.

A formal evaluation of the system was conducted by installing an ImmersaDesk and stereo monitor in a school. Second grade students participated in what was essentially the same regimen as the third pilot study. Performance of the VR group on a test of their knowledge of the round earth significantly improved over their pre-VR performance. In addition, student performance was significantly improved on a delayed post-test four months later. However, performance of a group that took the pretest and post-test without interacting with the virtual environment in between also significantly improved, so it’s hard to ascertain if the virtual environment really improved learning or not.

The Round Earth project sounds a cautionary note to those who would advocate blindly applying any new technology to the school system. Although they did find evidence that their system enhanced learning and retention of a
difficult concept, it was only after multiple design iterations. As the authors point out, student enjoyment of the experience isn’t enough of a justification to argue for the inclusion of virtual environments in the schools. There also has to be demonstrated improvement in student educational success for VR to crack the already crowded curriculum calendar.

VPS: The Virtual Puget Sound project was intended to study the effectiveness of an immersive virtual environment as a tool for teaching complex scientific concepts and principles. The initial subject matter was the ecology of the Puget Sound system, combining data from demographic, atmospheric, hydrographic, and oceanographic databases along with models of the region to predict environmental phenomena. Students are able to view a complete tidal cycle from many different viewpoints, and to examine data such as water speed and direction and salinity at twelve different depths, and at various playback speeds. Students can sample the data at five different locations, and can view and manipulate a graphical representation of the tide levels.

Before entering the virtual environment, students were shown a videotape which oriented them to the sound, showed how the data was collected, and showed models being constructed from the data. Students were then trained in using the interface and tools of the system, and then given a problem to solve. Possible problems included where to place the outflow from a sewage treatment plant, and to find the most likely habitats of a voracious species of fish that had been observed in the sound. A pre-test, post-test and an interview are used to determine what the students learned.

A pilot study with roughly 20 students was conducted. Students spend 30 minutes in training, and 60 minutes problem solving for a total of 90 minutes in the virtual world. There was a slight, statistically significant increase in the

24
student understanding of the water phenomena of Puget Sound. However, the virtual environment also generated serious misconceptions. Twenty percent of students that initially correctly understood that salty water settled below fresh water "learned" the opposite concept from using the system. In addition, a large minority of the students found the water current arrows, their lengths, and colors, confusing, and had problems remembering what each meant. Students did gain an appreciation of the complex nature of water flow in the sound, but interpreted the observed phenomena in many different ways, ascribing cause and effect relationships to things that were only secondarily related. The pilot study concluded that

Our experiences to date lead us to conclude that generalizations about learning in IVEs are difficult to make because learners' experiences are so varied, so dependent on the specific character of 1) visual representations, 2) the complexity of the underlying phenomena, and 3) the kind of structured experiences learners can have in these worlds.

Clearly this kind of basic research is necessary to reveal trends in how individuals use IVEs and to establish "proof of concept" that individuals can learn in virtual worlds; however, in a practical sense, virtual world experiences are of such brief duration they are limited in what they can tell us about conditions that would maximize learning. These optimal conditions, of course, would include 1) more extended opportunities for individuals to use the IVE, 2) the opportunity to have sense-making dialogue with others in between these experiences, and 3) embedding the use of IVEs in longer term curricular frameworks that include hands-on activities, readings, reflection, and some direct instruction.37

Yet again, this study has shown that VR might or might not be a useful educational tool, depending on the subject matter and also the system implementation. The determination of its utility has to be done on a case by case basis, at least until enough cases are tested that patterns start to emerge.

**ImmersaDesk in the Classroom** While it is useful to study virtual environments in the laboratory, in order for them to impact public education to any extent, they need to eventually move out into the school systems. The utility of a virtual reality system in an actual school setting is under investigation by The Electronic
Visualization Lab (EVL) at the University of Illinois at Chicago. An ImmersaDesk was installed next to the media center at an elementary school in Chicago. Since the media center had its own full-time teacher, the plan was for the regular teacher to work with one or a few students using the virtual environment while the rest of the class worked on other work in the media center. The project was composed of three thrusts: cognitive studies, quick worlds, and virtual ambients. In the cognitive studies, students did the usual pre-test, VR exploration, post-test. However, this approach was disruptive of the normal classroom routine, since students would be pulled out of class, individually, multiple times. In the quick worlds, VR environments would be quickly constructed at a teacher’s request, focusing on the parts of the world the teacher felt needed emphasized. Students would explore the environment in small groups under the guidance of the teacher for fifteen to twenty minutes. The objective of this thrust was to make the virtual environment just another tool, like a film strip or an interactive classroom activity. In virtual ambients, the entire class would explore a virtual environment in order to collect data, which would then be used to create and test an hypothesis. Rather than teaching a particular content area, the goal was to teach students the scientific method, and let them actually practice data collection instead of staring with pre-collected data.

Preliminary results from this work held several surprises. Even though the virtual reality equipment was in the school instead of in a university laboratory, it was still not convenient enough for teachers to use readily. It turned out that the media center had a tightly focused curriculum and that instead of some students exploring a virtual environment while the rest of the class did media center activities, the logistics of the situation caused students to be pulled out of regular class activities in order to participate in the virtual environments.

26
Therefore, the EVL has installed a simplified VR system (without head tracking, or stereo) in an actual fourth grade classroom, where it will be running all the time, and instantly available for use at any time during the day.

Students still found the chance to explore a virtual environment exciting, and the technology impressive. This interest helped keep the students motivated as they performed the learning activities. However, students are getting inured to high-quality computer technology as video game systems improve, and one frequently asked question about the virtual reality equipment is “What other games does it play? What else does it do?” The high interest students have because of the technology is paradoxically what has to be overcome if VR is to become a standard classroom technology, because if it comes into widespread use, the novelty will be gone.

One other result from their initial studies was to actually reduce the use of technology in their virtual environments. Initially, they always used head tracking so that the view would be stereoscopically correct for the person controlling the interaction. However, students had a tendency to make rapid head movements, which caused other non-tracked students to become ill, and when the tracker was switched from one student to another, the view tended to jump around jerkily, again causing discomfort in some students. Head tracking was actually disabled in the environments unless 1) a single student was working in the environment, or 2) a teacher was controlling the environment (since teachers knew the need to make slow head movements at all times). Thus the sense of immersion, and any enhancements to learning it produced were disabled in the interests of safety and student comfort.

The final lesson learned from this study was the difficulty in assessing the utility of a virtual environment in the classroom. There were so many factors that
might have contributed to the improvement or degradation of learning when comparing the traditional environment with the use of virtual reality to augment instruction that it was difficult for the EVL to determine whether or not the use of a virtual environment itself was making a significant difference.

Virtual Animals

The problem of constructing virtual models of animals is a fascinating one, and one that many researchers have tackled. The investigations run the gamut from trying to build robots that act like their animal counterparts to building animated computer animals that range from cartoon-like to quasi-realistic, to animal models intended for biological research. Some models run in real time while others take hours to simulate a minute in real life; some models are concerned with individual limb movements of individual animals while others focus on population trends of the whole. In order to build virtual gorillas, it was necessary to build individual animal models that would react in real-time in the way actual gorillas would, and to make movements that mimicked those of actual gorillas.

The following are representative of animal models that are concerned with accurately modeling animal motions and behaviors in real-time, or that are concerned with responding to unscripted user input.

Silas the Dog. Forming the basis for Bruce Blumberg’s Ph.D. research, Silas the Dog used a camera to capture video of the user and superimposed it in a fantasy world along with an animated dog^38. The dog, Silas, responded to the user’s actions and also initiated interactions on its own. The behavior controller used in Silas consisted of five levels arranged in a top-down fashion. There were three major levels and two interface levels between the major levels. The geometry level specified the parts of the creature model, while the degrees of freedom
(DOF) interface indicated what degrees of freedom were available and which were already being controlled. The motor skill level specifies the basic movements available to the creature. Any motor skills that were no longer being actively invoked would generally begin to return their various DOFs to a neutral position. The controller interface layer mapped general commands such as move forward into low level DOF changes necessary to cycle the geometry in order to produce a walking motion. Finally, the behavior level sensed the environment and selected the current creature behavior.

Behaviors ranged from low level (avoid walking into that wall in front of you) up through high level (satisfy you increasing hunger). The environment was sensed through the use of synthetic vision. In this process, the world was rendered from the view of the creature, and the resulting image was interpreted to provide information about the environment. In addition, internal state variables indicated such features as hunger, fatigue, boredom, and so on. These were combined in the behavior level and a single behavior was selected using a winner take all mechanism. Losing behaviors did have a chance to suggest modifiers to the winning behavior, and each behavior was scaled by an interest level value that was reduced as a behavior was performed and increased when it was no longer being done. As each behavior became more important, it simultaneously acted to inhibit all other competing behaviors. This process was iterated until only a single behavior was active, and that was the behavior selected to be performed in the next cycle.

Silas implemented 24 motor skills and 70 low level motor commands. He had 40 behaviors and 8 internal variables. He used a hierarchical control system to specify his actions at any point in time. However, his behavior control system

29
was lumped together into a single layer that handled everything from obstacle
avoidance to playing fetch with the user. By lumping all behaviors into a single
level, it would be difficult to implement the wide range of behaviors shown by a
pack of social creatures with all the nuances and gradations contained in their
elaborate social conventions.

**Alpha Wolf**  The Synthetic Characters Group at the MIT Media Lab developed
the Alpha Wolf project\(^{49,51}\) to study ways for synthetic characters to learn dynamic
social hierarchies. Initially shown at Siggraph 2001, the project allowed three
users to assume the role of three wolf cubs in a world populated by the three
cubs and three adult wolves. The rest of the world contained only a few trees,
both to simulate the feel of arctic tundra and because earlier projects had found
that having objects in the world that were not crucial to the concepts being
explored and with which users could not interact actually detracted from the
user’s interaction with the environment.

Users guide the actions of their wolf cub through sounds that they make.
Each user growls, howls, whines or barks into his microphone to direct his cub,
shaping its behavior within the context of previous social interactions.
Relationships to the other wolves are indicated by icons surrounding a user’s
screen, showing relative dominance and submission.

Interactions with the system take place over three different time scales.
In the shortest (less than ten seconds), the wolves use a dynamic dominance
value to modify the selected motion to reflect the current dominance
relationships. During the intermediate term (around a minute), the wolves
develop more fixed dominance relationships, remembering the results of
previous interactions and using them to set the tone for future ones. Finally, over
the long term (five to ten minutes), the cubs grow through adolescence into adults which form the core of adults that the next set of users interact with as cubs.

The alpha wolf system uses adaptation and the learning of social relationships to keep the system interesting and interacting with users in different ways within the guidelines of proper wolf social behavior. If a system exhibits repetitive behavior without evolving, users quickly grow bored of interacting with it, and their motivation to continue interacting (and learning, if the system is educational) is greatly reduced. By allowing each set of users to determine how the dominance hierarchy is established for their particular set of cubs, each experience is different.

One problem with allowing multiple users in an environment is that it affects the balance between how much of the behavior is determined by the users and how much by the control system. If the users interact with each other through their avatars in ways unlike the animals they are supposed to be controlling, then the animals behave unnaturally. For an educational system, this would mean users learning incorrect behaviors for the animal under study. However, if the users’ creatures are overly controlled by the computer, then the element of interactivity is lost and the system becomes nothing more than a badly done video. Without some element of control, users are not free to explore and learn at their own pace. The alpha wolf system allows social relationships between the wolves to be established by the user interactions, but no mention is made of what would happen if all three users consistently tried to become the most dominant wolf.
Fish: A landmark simulation of fish was constructed by Tu and Terzopoulos. Fish motions were generated and controlled with a mass-spring system, which simulated the skeleton and muscles of the fish. Specifying changing rest lengths for the springs would cause the fish to move its tail from side to side. A hydrodynamic model would compute the interaction of this motion with the surrounding water and calculate the resulting translation of the fish. A later version of the system generated motion using a genetic algorithm to optimize the distance traveled while minimizing energy use. The fish sensed their environment through a temperature sensor which sensed the surrounding water temperature, and a vision sensor which senses in a 360 degree spherical angle.

The fish had three mental state variables measuring hunger, libido, and fear. An intention generator used the mental state, incoming sensory information, and the fish’s habits to generate an intention which selected and executed a behavior routine, that then drove the motor controllers. The fish had eight behavior routines to do such things as schooling, wandering, feeding, avoiding obstacles, and so forth. The fish intention generator also had a single item of short term memory to remember the current behavior in case it needs to be temporarily overridden by a higher priority behavior. This was used to reduce dithering among fish behaviors.

Avoiding collisions was the highest priority of the behavior controller. If no collisions were imminent, the next behavior considered depended on whether the fish was a predator, prey, or pelagic fish. Predators would wander or eat depending on how hungry they were, while prey would school or flee if a predator
was near, or else eat or mate depending on hunger. Pacifists did not need to worry about predators and concentrated on mating and eating.

Behavior control was done using a decision tree. Higher priority behaviors were considered first, and if they were not necessary, the next lower priority of behavior were considered next. Since there were only eight fish behaviors to be dealt with and they could be arranged in priority order, the behavior controller did not need to be overly complex. Because of this, the focus of this project was on realistic motion generation. Also, there was no provision for the user to interact with the fish. Users could only observe the simulation as the fish interacted with each other.

**Reynolds’ Group Behaviors.** Craig Reynolds’ work on flocking and herding behaviors of birds, fish, and other creatures has been the basis for a number of movie special effects, from the spiders in *Arachnophobia* to the wildebeests in *Lion King*. A simple system of three rules resulted in reasonable looking flocking behavior:

1. avoid collisions with nearby flock mates
2. attempt to match velocity with nearby flock mates
3. attempt to stay close to nearby flock mates.

Lower numbered rules take precedence over higher numbered ones, so collision avoidance would take priority over flock centering. His work showed that complex, interesting behaviors could come from a few simple rules. Also, his system of having a linearly ordered list of rule priorities where lower numbered rules override higher numbered rules proved useful in the design of the three level gorilla controller, in which lower level actions and reactions preempted higher level ones.
CHAPTER 3
BUILDING THE VIRTUAL GORILLA ENVIRONMENT

The Virtual Gorilla Environment was designed to provide information to students in five areas: gorilla visual appearance and identification, gorilla vocalizations, zoo habitat design, the gorilla dominance hierarchy, and gorilla social interactions. These were all areas where Zoo Atlanta’s exhibits were not as effective as they needed to be for the zoo to accomplish its educational mission.

Gorilla Identification

Although each exhibit at Zoo Atlanta had signs posted at the viewing area listing the gorillas in the exhibit, and providing information such as their weights, names, and types, it was difficult for zoo patrons to match the names to the actual gorillas. They would confuse males and females, juveniles and adults, and generally have a hard time deciding which gorilla was which. One educational objective of the Virtual Gorilla Environment was to provide enough information and training that users could more easily identify the different types of gorillas. While a user might still confuse two females, for instance, at least after exploring the environment the user should be able to properly identify silverbacks, females, and juvenile gorillas, which were the initial types inhabiting the virtual environment.

In building the virtual gorilla models, there were two competing considerations. On the one hand, for students to be able to effectively identify the different gorilla types, it was necessary that the gorilla models approximate the appearance of real gorilla types. On the other hand, making the models too detailed would increase the number of polygons and/or textures needed for each gorilla, which would have a negative impact on the frame rate of the system. In any VR system, one of the paramount concerns is

34
the issue of frame rate. If the frame rate was too low, the resulting lag destroys the sense of immersion, and could make the user physically ill. Therefore there was an impetus to use as few polygons and textures as possible, especially since the initial development system (an SGI Crimson Reality Engine) had a texture memory limit of 1MB—beyond that, textures had to be paged in and out of memory with a resulting degradation of frame rate. Empirical determination from other environments indicated that the initial hardware system could support approximately 10000 polygons in the environment and maintain at least a six frames a second update rate. The goal was a minimum of 10 frames per second. But other environments running as slow as six frames per second had proven useful, so this was the absolute minimum that would be accepted for the virtual gorilla environment.

After building a model of the gorilla habitat and ruthlessly simplifying it, there were enough unused polygons to use gorilla models with between 2000 and 3000 polygons each and to still have two or three virtual gorillas in the environment.

Special care was taken when modeling the gorillas to ensure that body proportions used were scaled to match the environment. In this way, students could get a feel for how big a real gorilla was and what its body proportions were, without drawing erroneous conclusions from incorrect size measurements. Jungers provided a formula for calculating approximate limb lengths (upper and lower arm, upper and lower leg) based on gorilla mass:

\[ y = bx^k \]

where \( y \) is the limb length in millimeters, \( x \) is the mass of the gorilla in grams, and \( b \) and \( k \) are constants that depend on the limb length being specified. Jungers reported the typical mass of a male gorilla as being 169.5 kg for a gorilla gorilla gorilla, based on a
study of 21 adult skeletons, and 159.2 kg for a gorilla gorilla beringei, based on 7 skeletons. Similarly, masses for a female from 18 skeletons were 71.5 kg for a gorilla gorilla gorilla, and, based on 8 skeletons, 97.7 kg for a gorilla gorilla beringei. For African apes, the constants were determined to be as shown in Table 3-1.

<table>
<thead>
<tr>
<th>Body Part</th>
<th>b</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humerus</td>
<td>16.46</td>
<td>0.272</td>
</tr>
<tr>
<td>Radius</td>
<td>39.23</td>
<td>0.181</td>
</tr>
<tr>
<td>Femur</td>
<td>55.95</td>
<td>0.156</td>
</tr>
<tr>
<td>Tibia</td>
<td>56.58</td>
<td>0.137</td>
</tr>
<tr>
<td>Forelimb</td>
<td>48.75</td>
<td>0.230</td>
</tr>
<tr>
<td>(Humerus + Radius)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hindlimb</td>
<td>111.97</td>
<td>0.148</td>
</tr>
</tbody>
</table>

Using a mass of 165 kg for a typical male gorilla and a mass of 80 kg for a typical female, the limb lengths shown in Table 3-2 were derived using the formula.

<table>
<thead>
<tr>
<th>Body Part</th>
<th>Male Length (in mm)</th>
<th>Female Length (in mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humerus</td>
<td>432.1</td>
<td>354.9</td>
</tr>
<tr>
<td>Radius</td>
<td>341.5</td>
<td>302.7</td>
</tr>
<tr>
<td>Femur</td>
<td>364.5</td>
<td>325.6</td>
</tr>
<tr>
<td>Tibia</td>
<td>235.5</td>
<td>267.6</td>
</tr>
<tr>
<td>Forelimb</td>
<td>772.7</td>
<td>654.2</td>
</tr>
<tr>
<td>Hindlimb</td>
<td>662.7</td>
<td>595.3</td>
</tr>
<tr>
<td>(should be same as Forelimb)</td>
<td>777.2</td>
<td>667.6</td>
</tr>
<tr>
<td>Femur + Tibia (should be same as Hindlimb)</td>
<td>669.0</td>
<td>593.2</td>
</tr>
</tbody>
</table>

As can be seen from Table 3-2, the values given by the formula for the parts of the forelimb and hindlimb sum up fairly closely to those computed from the formula for the forelimb and hindlimb.
Jungers also reported indices for comparing limb proportions between members of closely related taxa. Using the formulas and constants given for gorilla gorilla, based on 21 males (Table 3-3), a second set of lengths was derived. These are shown in Table 3-4. Combining the two sets of values resulted in the limb lengths shown in Table 3-5, which were the lengths used to model the silverback. Because the various indices were based only on male gorilla measurements, female limb lengths shown in Table 3-2 were used when modeling the female gorilla.

Limb circumferences were difficult to find in the literature, and evidently the Yerkes Primate Research Center considered the information on their gorillas (which made up the largest part of the exhibit at Zoo Atlanta) to be proprietary, because it proved impossible to obtain the measurements from Zoo Atlanta. Finally Kyle Burks was able to provide circumferential data for some of the limbs of various gorilla types, which are shown in Table 3-6. These measurements were used to scale the virtual gorilla limbs by placing them in a cylinder of the specified diameter, aligning the limb with the cylinder axis, and scaling the limb so that it was just tangent with the inside wall of the cylinder.

Table 3-3. Gorilla Ratios of Limb Lengths (Based on 21 Adult Males)

<table>
<thead>
<tr>
<th>Index</th>
<th>Definition</th>
<th>Male Gorilla Value</th>
</tr>
</thead>
</table>
| Intermembral Index         | \[
\frac{\text{Humeral + Radii}}{\text{Femur + Tibia}} \times 100
\] | 115.6               |
| Humerofemoral Index        | \[
\frac{\text{Humeral}}{\text{Femur}} \times 100
\] | 117.2               |
| Brachial Index             | \[
\frac{\text{Radius}}{\text{Humeral}} \times 100
\] | 80.4                |
| Crural Index               | \[
\frac{\text{Femur}}{\text{Tibia}} \times 100
\] | 62.9                |

\[\text{Forelimb length (mm)} = \frac{\text{body weight (g)}}{30} \]

\[\text{Hindlimb length (mm)} = \frac{\text{body weight (g)}}{25} \]

37
Table 3-4. Male Gorilla Limb Lengths from Limb Proportions

<table>
<thead>
<tr>
<th>Body Part</th>
<th>Length (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humerus</td>
<td>441.9</td>
</tr>
<tr>
<td>Radius</td>
<td>355.3</td>
</tr>
<tr>
<td>Femur</td>
<td>377.4</td>
</tr>
<tr>
<td>Tibia</td>
<td>312.8</td>
</tr>
<tr>
<td>Forelimb</td>
<td>690.2</td>
</tr>
<tr>
<td>Hindlimb</td>
<td>797.2</td>
</tr>
</tbody>
</table>

Table 3-5. Working Values for Male Gorilla Gorilla Limb Lengths

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>170 kg</td>
</tr>
<tr>
<td>Humerus</td>
<td>0.442 m</td>
</tr>
<tr>
<td>Radius</td>
<td>0.355 m</td>
</tr>
<tr>
<td>Femur</td>
<td>0.377 m</td>
</tr>
<tr>
<td>Tibia</td>
<td>0.313 m</td>
</tr>
<tr>
<td>Forelimb</td>
<td>0.797 m</td>
</tr>
<tr>
<td>Hindlimb</td>
<td>0.690 m</td>
</tr>
</tbody>
</table>

Table 3-6. Gorilla Limb Circumferences

<table>
<thead>
<tr>
<th>Gorilla Type</th>
<th>Upper Arm</th>
<th>Lower Arm</th>
<th>Thigh</th>
<th>Calf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult Male</td>
<td>45</td>
<td>40</td>
<td>64</td>
<td>35</td>
</tr>
<tr>
<td>Adult Female</td>
<td>33</td>
<td>27</td>
<td>48</td>
<td>29</td>
</tr>
<tr>
<td>Juvenile</td>
<td>27</td>
<td>26</td>
<td>44</td>
<td>27</td>
</tr>
</tbody>
</table>

For gorilla young, Jungers' formulae were not guaranteed to be valid.

Measurements for three different gorilla young were obtained from Fossey’s, and are given in Table 3-7.
Table 3-7. Body Measurements for Young Gorillas

<table>
<thead>
<tr>
<th></th>
<th>3 Month Old Female</th>
<th>39 Month Old Male</th>
<th>46 Month Old Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humerus</td>
<td>127.5 mm</td>
<td>255 mm</td>
<td>210 mm</td>
</tr>
<tr>
<td>Radius</td>
<td>102.5 mm</td>
<td>220 mm</td>
<td>210 mm</td>
</tr>
<tr>
<td>Femur</td>
<td>92.5 mm</td>
<td>130 mm</td>
<td>180 mm</td>
</tr>
<tr>
<td>Tibia</td>
<td>120 mm</td>
<td>250 mm</td>
<td>165 mm</td>
</tr>
<tr>
<td>Torso Height</td>
<td>190 mm</td>
<td>460 mm</td>
<td>490 mm</td>
</tr>
<tr>
<td>Head Height</td>
<td>100 mm</td>
<td>145 mm</td>
<td>765 mm</td>
</tr>
<tr>
<td>Head Width</td>
<td>80 mm</td>
<td>130 mm</td>
<td>125 mm</td>
</tr>
<tr>
<td>Hand Length</td>
<td>80 mm</td>
<td>155 mm</td>
<td>150 mm</td>
</tr>
<tr>
<td>Foot Length</td>
<td>102.5 mm</td>
<td>172 mm</td>
<td>167.5 mm</td>
</tr>
</tbody>
</table>

The models generated were simplified as much as possible while maintaining the correct body segment sizes. The result was a gorilla family whose members are composed of between 2000 and 3000 polygons each, which conveyed the size and shape of each member without too severe a degradation in rendering performance. Each member was colored a grayish-black approximating the color of gorilla fur. Although many attempts were made to find and apply a reasonable fur texture, they were uniformly unsuccessful, so no fur textures were used. The difficulty was caused by a combined problem of finding a fur texture that would tile seamlessly, while also finding one that approximated the interaction of fur with sunlight. In addition to size being a distinguishing feature between the silverback and the female models, the back and hindquarters of the silverback were colored silver, to represent the silver colored fur of the silverback. For the female model, appropriately distorted hemispheres were added to the chest to represent mammary glands, which are visually prominent on female gorillas. A comparison of the results to actual gorillas can be seen in Figures 3-1 and 3-2.

Each gorilla model was specified using configuration files that were read when the program was run. In this way, new or improved models could be incorporated into
the program without having to recompile the source code. Initially, the models were composed of 9 body parts, 8 joints, and 14 degrees of freedom. This proved inadequate for the range of motions that needed to be executed, so the models were re-specified to have 11 joints and 28 degrees of freedom. Each body segment had its inner joint centered at the origin in its local coordinate system, and was then translated to the appropriate place relative to its parent object. This allowed motions to be specified using relative joint angles, making the specification more general than would otherwise be possible.

![Modeled Silverback](image1)

![Actual Silverback](image2)

Figure 3-1. Modeled vs Actual Silverback

![Modeled Female](image3)

![Actual Female](image4)

Figure 3-2. Modeled vs Actual Female
Making the virtual gorillas as representative as possible was just the first step, though. Just because a student noticed that gorillas with light colored bottoms seemed to be in charge did not mean that they would recognize the dominant gorilla as a male silverback. To make sure that that connection was made, the system included audio annotations naming the various gorilla types as the user looked at them, and indicating their current moods. In this way, as the user looked at or interacted with the different virtual gorillas, he would have their identifications repeatedly associated with their body features, facilitating learning to identify the different gorilla types.

Gorilla Vocalizations

Teaching users about the sounds and meanings of various gorilla vocalizations was seen early on as a potential advantage that a virtual environment might have over the actual gorilla habitats at Zoo Atlanta. While the habitats at Zoo Atlanta allowed visitors to observe gorillas in reasonably lifelike settings, the gorillas were still far enough away from viewers that it was difficult to hear any vocalizations they made. The one place where visitors were just a few feet from the gorillas was the visitors' center, and in this case there was a thick sheet of bolted glass separating the gorillas from their observers, which effectively blocked all sounds (except for the gorillas barking on the glass, of course). However, vocalizations play an important role in gorilla social interactions, and so they were deemed an important educational objective of the virtual gorilla environment.

Because of their collaboration with the producers of the movie *Congo*, Zoo Atlanta was given a professionally recorded CD of the vocalizations of the actual gorillas at the zoo. A copy of this was made available to generate the vocalizations of the virtual gorillas. These were separated into three categories as representative of the vocalizations of gorillas: contentment vocalizations, vocalizations of annoyance, and
expressions of anger. Contentment vocalizations of the various inhabitants of the zoo were combined into two sequences, one for females and one for silverbacks. These were looped with appropriate pauses, and played whenever a gorilla was content. Each loop was between one and two minutes in length, and displayed several different rumbles to the user. Two samples of warning coughs were provided, one male and one female, which were played whenever one of the virtual gorillas was annoyed with another, or with the user. Finally, a recording of chest beating and a roar was used as an anger sound by all of the virtual gorillas. Each virtual gorilla would play back the appropriate sounds indicative of its own mood. The sounds would also decrease in intensity the farther away the user was from the corresponding gorilla, so that sounds from a gorilla the user was interacting with would dominate the soundscape.

One of the gorilla habitats had a stream that ran down into one of the moats, providing a background sound of water splashing. A recording of a babbling brook was looped to recreate this, and to provide a background of white noise to help mask out external sounds.

A gorilla's mood and current interaction state would drive which motion and which keyframes would be displayed next, and these keyframes would trigger their corresponding vocalizations. Thus, vocalizations reflecting a gorilla's mood would be played whenever appropriate. By observing a gorilla's interactions and corresponding vocalizations, a user should be able to deduce the mood and meaning reflected in each vocalization.

Zoo Habitat Design

A lot of thought goes into a zoo habitat design, and there are conflicting opinions as to the best design. In addition, the design has to work with whatever terrain is available at the zoo, which imposes further restrictions on it. Since it is a major
undertaking to change a habitat design, as much as possible it must be done correctly the first time. However, when done correctly, many of the design issues are subtle enough that they are missed by observers, and the information about an animal’s native habitat that is to be conveyed is overlooked. Thus, the design of the gorilla habitats was something that was of interest educationally, but often overlooked by students just observing gorillas in the habitats.

Since a virtual environment modeled after Zoo Atlanta’s habitats would have some of the same drawbacks as the actual habitats did, an additional feature was added to the virtual gorilla system to help highlight this information. As the user approached a significant feature of the habitat, an audio annotation would play, describing the feature and its importance in the design of the habitat. Annotations were included to explain that the masts kept gorillas separated from each other and from the humans while not inhibiting the views each had of the others as fences would. Additional annotations explained that the rocks had been included in the environment to provide a high place from which gorillas might display to other gorillas in their own or in other exhibit areas, and that the dead trees were provided for the gorillas to have something to rest against and with which to play. A complete list of all the audio annotations used, including those describing the habitat, identifying the various gorilla types, explaining the social interactions and gorilla moods, and discussing the gorilla dominance hierarchy are listed in appendix B.

In addition to the audio annotation explanations, by allowing the users to actually explore the entire habitat, they would be able to discover features not visible from the periphery to which visitors were normally restricted. This included such features as areas in the viewable habitats not visible from any of the viewing locations where a
gorilla could go to get away from staring people, and even an entire habitat that was out of view, and used for gorilla introductions or to isolate ailing gorillas.

![Modeled Habitat](image1) ![Actual Habitat](image2)

**Figure 3-3. Modeled vs Actual Habitat**

**Gorilla Dominance Hierarchy**

Gorillas have a dominance hierarchy which is reflected in their social interactions. Although it is not always constant, in small groups of gorillas it can be approximated by a linear hierarchy, with the silverbacks as the most dominant, then the blackback males, followed by the females, juveniles and infants. In reality, depending on the standing of the father and mother, infants are excused for taking liberties that are not permitted older gorillas. In addition, if the dominant female has been with the group for a long time, she might be more dominant than some of the younger blackback males. However, as a first approximation, a linear dominance hierarchy is a good representation, and can be used to illustrate the way a dominance hierarchy controls social interactions.

The virtual gorilla environment used a linear dominance hierarchy to control interactions among virtual gorillas, and between virtual gorillas and the user. A disruptive user who repeatedly defied the social conventions was removed from the environment and placed is time out, and then restarted in the model of the visitors' center with an audio annotation explaining the dominance hierarchy. In addition, mood
indicators were provided over each gorilla's head so that the user could determine which gorillas were content, which were annoyed, which were angry, and which were submissive. By running around the environment and interacting with the virtual gorillas, the user could discover the place of a juvenile gorilla, represented by the user, in the dominance hierarchy. In addition, by observing interactions among the virtual gorillas, especially ones instigated by having one virtual gorilla chase the user into a group of virtual gorillas, the user could examine interactions among different types of virtual gorillas in the environment. Depending on the speed of the graphics rendering, there was a virtual silverback as head of the family group, and then two or three females. Occasionally there was another juvenile gorilla in the group as well, but in every case, all the virtual gorillas were higher in the dominance hierarchy than the user.

By making the user the lowest member of the dominance hierarchy, the user could learn about it just as any juvenile gorilla would, through trying things and being put in his place. Since as infants grow into juvenile status they are expected to start behaving appropriately and their incorrect or deviant behavior is lovingly punished, this seemed a good place to start students who wouldn't necessarily know correct gorilla behaviors and might act in a similarly strange fashion. Starting users as more dominant gorillas would put them in a role where they would be expected to know how to behave appropriately towards more submissive group members, and it would be much harder or impossible for them to discover correct behavior, since the more submissive gorillas would accept any behavior, no matter how bizarre.

Gorilla Social Interactions

One of the most important educational objectives of the virtual gorilla environment was to teach students about gorilla social interactions. Gorillas are social animals, and use the dominance hierarchy to control a fairly stylized set of social
interactions. Each gorilla type spends differing amounts of time in affiliative, agonistic, and solitary behaviors, but each does spend some time in each behavior. Affiliative behaviors include invitation to play and play, and invitation to groom and grooming. Agonistic behaviors include assertion or reinforcement of the dominance hierarchy dealing with issues of proximity or gaze, and challenging for dominance. Solitary behaviors are behaviors such as foraging for food, object manipulation, sitting, sleeping, and so on.

While solitary and affiliative behaviors can be observed when watching gorillas on exhibit at the zoo, agonistic behaviors are seldom seen. By the time that gorillas are allowed on exhibit, they have already determined their positions and social standing in the group in private, and thus only occasionally need to have the hierarchy reinforced. While allowing a gorilla group to work out its rankings in a controlled environment and without the additional stress of hundreds of zoo visitors staring at them, zoo visitors also seldom realize the existence of a dominance hierarchy since the gorillas on exhibit all know their places, and for the most part, remain in them! For this reason, the virtual gorilla environment focused on implementing the dominance hierarchy and agonistic social behaviors.

**Gorilla Motions.** In order to implement representative interactions, motions used during each interaction needed to be scripted for each gorilla type. Also, a basic repertoire of motions indicative of solitary behavior were needed for the virtual gorillas to use when they were not interacting with each other. The motions generated were based on those made by the actual gorillas at Zoo Atlanta as observed and recorded on video tape. Additional motions were generated based on explanations and demonstrations by the gorilla research staff at the zoo. These motions were stored as a sequence of
timestamped keyframes, and linear interpolation was used between keyframes to 
generate intermediate poses.

Each keyframe was stored using relative joint angles. The root of the kinematic 
chain was the center of mass of the torso, and this had six degrees of freedom, allowing 
the gorilla to be positioned anywhere in the habitable model, facing any direction, and with 
any orientation. Terrain following was implemented by using interpolation of a sampled 
heightfield map to determine the elevation at the hands and feet and to adjust the center 
of mass upward or downward appropriately in order to maintain contact without 
penetration of the surface. The height field also encoded forbidden regions where either 
the virtual gorillas or the user were not allowed to go, and was used to enforce that 
constraint.

Transition between sequences of keyframes was a very important part of 
producing realistic-looking motion. There are certain motions that flow fluidly from one to 
the other, and others that make no sense when interpolated between. For instance, 
when interpolating directly between a gorilla lying on its left side and one standing on all 
fours, the gorilla would appear to rotate while levitating vertically before finally standing, 
and in the process assume some physically unrealistic postures. On the other hand, the 
transition from sitting to standing on all four feet looked smooth and natural.

The solution chosen here was to generate a table of allowed motion transitions. 
These were used to specify what motions were allowed given the current motion in 
progress. For instance, if a gorilla were lying on its left side and wished to stand on all 
fours, it would select standing upright as the next desired motion. Looking this up in the 
transition table, it would discover that it was impossible to go directly from lying on the 
left to standing on all fours. However, the table did more than merely indicate whether a 
transition was allowed or not; it also suggested the best next motion to choose based on
the desired motion. In the example of switching from lying on the left, the transition table would suggest choosing a sitting position as the next motion. The system would then interpolate between lying on the left and sitting.

Once the gorilla was sitting, assuming that the desired motion was still standing on all fours, the transition table would indicate that this was an acceptable transition from the sitting position, and the transition would be made. The end result would be a motion sequence in which the gorilla would first sit up, and then stand up, a much more realistic-looking sequence than the one progressing directly from lying to standing.

As implemented to this point, the system would have a set of motions representing various behaviors that could be used when gorillas interacted with each other or the user, which would illustrate the dominance hierarchy. In fact, an early version of the system was implemented this way, with a knowledgeable gorilla expert directing the virtual gorilla behaviors behind the scenes. For the system to be able to run without the guidance of such an expert, though, it was necessary to implement behavior controllers for the virtual gorillas which would choose the appropriate motions to represent the desired behaviors and reflect the current moods of the gorillas.

Gorilla Behavior Controller In order to program lifelike gorilla behaviors and social interactions, the behaviors and interactions had to be quantified and encoded as heuristics. This proved to be amazingly difficult, even for the limited set of behaviors that were chosen to be implemented. It was necessary to determine the size and shape of the personal space of each type of gorilla, to rank the gorillas in a dominance hierarchy (which was assumed to be linear as a simplifying first-order approximation, generally true for small groups of gorillas but not necessarily so for larger groups, or groups with infants), to determine a range of situations that would be modeled between the user and a gorilla and among gorillas, and to determine and describe quantitatively the behavior
of each type of gorilla in each type of situation. Since any gorilla type could interact with any other, and since there were five basic gorilla types, there were twenty five possible sets of interaction behaviors for two gorillas. For larger groups of gorillas all interacting with each other, the number grows exponentially. Clearly, the behavior modeling task could have gotten out of hand quickly.

Building accurate models of animal behaviors was both easier and harder than building accurate models of human behavior patterns. On the one hand, as a human, the system architect could use introspection to determine underlying causes for behaviors, enabling generalizations to be made that simplify system design. For other animals, while ethologists can speculate as to their internal motivations, all that the system architect had to reliably build upon was observations of external behaviors.

On the other hand, animal behaviors appear to be much less complex than human behaviors, so that the behavior control mechanism could be correspondingly simpler, having to deal with fewer, simpler situations.

Since it was possible only to infer internal state or motivations for non-human creatures, a behaviorist approach to the design of action selection mechanisms for virtual animals was taken. Even if behaviorism turned out to be an incorrect or incomplete scientific theory, it did provide a reasonable basis for building virtual animal behavior controllers, since all the system architect has to work with reliably was observed external behaviors generated as reactions to varying stimuli, as reported in the literature.

The biological literature on gorillas was used to generate a set of simplifying assumptions, which were then reviewed and modified by the gorilla experts at Zoo Atlanta. These were then implemented in a parameterized behavior controller for each gorilla type. The basic interactions were programmed once, and the parameters were used to customize the interactions for each gorilla type. Each gorilla type had a
behavior control file that specified the sense routine, action selection routine, movement routine, and pose routine to be used for this type, using a table driven dispatch system. In addition, each behavior control file specified such parameters as how large that type’s personal space was to the front, sides, and rear, how close one had to be before staring was considered rude, and the lengths of time spent in each of the various mood states before transitioning to another state. The actual values used are documented in Appendix C.

The behavior control architecture that seemed the most logical for any type of virtual animal, and that was used when building the virtual gorillas was to take a layered approach, building from the bottom up, in a manner similar to Brooks’ subsumption architecture. However, unlike Brooks’ systems which had a potentially infinite number of interacting layers, it turned out that most animal behaviors fell into three distinct types. Each of these types could be built as a layer, with the higher level behaviors built on top of the lower layer(s). The three types were the reflexive, the reactive, and the reflective behaviors (see figure 3-4). In general, behaviors at a lower level had priority over behaviors at a higher level, and preempted them.
At the lowest level, the reflexive behaviors included such actions as avoiding obstacles. In the virtual gorilla system, these included the interpretive center, the moats, and the rocks. These actions were taken in situations that demanded an immediate response, and took priority over any other action. Even when fleeing from a predator, for example, an animal would avoid running into a rock or tree as it fled, and placing this type of behavior at the lowest level allowed the virtual gorillas to behave similarly.

In the virtual gorilla system, creatures sense objects ahead of them as they are moving, and turn away, turning away at a greater rate the closer the obstacle is to them. Because obstacles include the moats which form an irregular boundary around the entire environment, simple distance calculations to obstacles are not used, as would be possible if only trees or rocks were to be avoided. Instead, various spots in front of and to either side of the creature are sampled, and orientation modified based on the results. This occasionally proved to be a less than optimal solution, but worked well enough for the prototype system.
At the next level, the reactive behaviors were those dealing with interactions among creatures. For many animal societies, each animal's position in that society determines how the other members interact with that animal. Although other proposed behavior controllers have focused on predator-prey relationships between species, within a species it is the more subtle dominance-submission relationship that determines animals interactions towards each other. For the virtual gorillas, if they were not in immediate danger of running into an obstacle, their next concern was whether or not there was the possibility of any interaction with another gorilla. Gorillas have a fairly well defined personal space, and have rules about who can or cannot violate that personal space. They also have taboos about staring at other, more dominant gorillas. Although in large groups the dominance hierarchy can be rather complex and vary depending on the situation, for small groups such as those at the zoo, a simple linear ordering was a reasonable approximation, and that was what was used in the virtual gorilla system.

Whenever a virtual gorilla was determining what to do next, it examined its environment to determine if there were other gorillas trying to interact with it, or if it had initiated an interaction with another gorilla. If so, then it used the dominance hierarchy to determine the appropriate behavior, and selected that action to perform next. Thus, if the silverback was chasing the student and the student ran toward a female, the female would stand and move out of the way since the silverback was more dominant. However, if the student ran toward a female without being chased by the silverback, the female would stand her ground and warn the student (who played the role of a juvenile, and so was lower in the dominance hierarchy) away with gaze aversion and warning coughs.

At the third layer of the behavior control architecture, activated only when not preempted by the two lower layers, was the reflective layer. This controlled solitary behavior, such as sitting and thinking or sleeping, solitary play, object manipulation,
feeding, and so on. Because the dominance hierarchy was preeminent for gorillas (as it is for many other species of animals), interactions with other gorillas took precedence over solitary actions. For example, if a less dominant gorilla was contentedly sitting in one spot and a more dominant gorilla approached, the submissive one would stand and move away, while the more dominant one sat where the less dominant one was sitting. Such displacement behavior is seen in real gorillas, and was a natural result of the behavior control architecture described here.

At the present, there is a paucity of reflective behaviors for the virtual gorillas. Currently they just sit or sleep—the problem has been the difficulty of implementing models of other motions. As the range of reflective behaviors becomes more complex, a more complex controller based on internal state (for example, hunger or fatigue) would be needed to select the appropriate reflective behavior. However, the basic three tier behavior control architecture should continue to cope satisfactorily with interactions between the layers.

One other problem present in many behavior controllers is that of perseverance. Pure reactive controllers are prone to dithering and to single goal fixation. For example if a creature was very hungry and moderately thirsty, many controllers would have it fixate on getting food, passing up the chance to opportunistically drink as it passed by a water source until the more dominant goal of reducing hunger had been attained. If the stronger goal was blocked (for instance no food is safely available), the creature would become stuck on that goal and not satisfy any of the lesser goals, even if they weren’t blocked. On the other hand if hunger and thirst were equally strong, simple behavior controllers often dithered back and forth, first starting to eat until hunger drops just below thirst, then starting to drink until thirst drops below hunger, over and over again. Real creatures will generally take advantage of opportunities to satisfy secondary needs on
the way to satisfying the primary one, and will also persevere in an activity, once started, until that particular need is sated.

As more behaviors are added to the virtual gorilla system in future versions, this potential problem needs to be kept in mind. In the current system, perseverance was handled by requiring cyclical actions such as walking to play out a complete cycle before being preempted. This way, if a less dominant gorilla stood up and walked off as a more dominant one approaches and sat, it would actually move completely out of the more dominant gorilla’s personal space. (Initially the less dominant one would walk until just outside the more dominant one’s personal space, and then immediately start to sit down—which would move it just over the boundary of the more dominant one’s personal space. This would result in a sequence of motions that looked like the less dominant gorilla was trying to scoot away while sitting on its bottom.)

At this point the system consisted of a reasonable model of Zoo Atlanta’s gorilla habitat three, along with virtual gorillas who would exhibit appropriate social interactions towards each other and the user based on the dominance hierarchy. Audio annotations had been included to explain factors that early testing showed were not intuitively obvious when exploring the environment without guidance, along with appropriate gorilla vocalizations reflecting each gorilla’s mood. All of the major educational objectives had been included in the system design. It was now time to test the effectiveness of the system. The next chapter will discuss the testing procedure and examine the results collected.
CHAPTER 4
TESTING THE VIRTUAL GORILLA ENVIRONMENT

Qualitative Testing

A prototype of the virtual gorilla environment was field tested at Zoo Atlanta, using students who had been participating in the zoo’s Young Scientist program. These students, from Westminster School, Trickum Middle School, Midway Elementary School, Slaton Elementary School, and Fayetteville High School in Atlanta, had been coming to the zoo weekly to learn to take behavioral data observations, and to use these observations to draw conclusions about gorilla behavior. Since these students were already accustomed to visiting the zoo and working with the gorilla exhibit staff, a version of the virtual gorilla environment was taken to Zoo Atlanta and set up in the Gorillas of Cameroon pavilion for the day. This setup used an SGI Onyx Reality Engine 2 to generate the graphics, and an SGI Indy to generate sounds. Head tracking was provided by a Polhemus Long Ranger, while video was presented to the user through a Virtual Research VR-4 head-mounted display. Users input movement commands using buttons on a joystick handle as they stood on a platform under which rested a subwoofer. Monitors were provided for other students to see what the user was seeing, so that they could make comments or suggestions to the user. Figure 4-1 presents two views of the prototype system being tested at the zoo.
Figure 4-1. Prototype System Test Setup at Zoo Atlanta

From 9:30am until 4:00 that afternoon, a steady stream of kids showed up to test the system. The reaction of the students that participated in testing that first prototype at the zoo was uniformly positive. Students stated that they thought it was fun, and that they felt like they had been a gorilla. More importantly, they appeared to learn about gorilla behaviors, interactions, and group hierarchies, as evidenced by their later reactions when approaching other gorillas. Initially they would just walk right up to the dominant silverback and ignore his warning coughs, and he would end up charging at them. Later in their interactions, though, they recognized the warning cough for what it was and backed off in a submissive manner. They also learned to approach the female slowly to initiate a grooming session, instead of racing up and getting bluff-charged. The observed interactions as they evolved over time gave qualitative support to the idea that immersive virtual environments could be used to assist students in constructing knowledge from a first-hand point of view.

Since each user was free to explore as he wished, with minimal guidance from one of the project staff, each could customize his VR experience to best integrate his new knowledge in terms of his pre-existing knowledge base. It was interesting to note that younger students spent more time exploring the environment, checking out the corners of the habitat and the moat and trying to look in the gorilla holding building. Older students spent more time observing and interacting with the other gorillas. Each tailored his
experience to his interests and level of maturity, yet everyone spent some time on each of the aspects (investigating the habitats, interacting with the other gorillas).

Also, even though students were free to interact with the environment in novel ways, most users interacted as they would have if they had actually physically been in the real environment. For example, the moats were 12 feet deep, and in the real world most people don’t willingly jump into 12 foot deep ditches. Even though the virtual environment was designed to allow users to easily enter and leave the moats, few did. Also, most users avoided running into the rocks on the habitat building wall, or trying to fly through trees, and had to be coaxed up to the top of the rocks in the gorilla yard initially. It seemed reasonable to infer from this that the students transferred their knowledge of the real world to the virtual one quite easily, and that their sense of immersion was good.

Prototype Revisions

Feedback from these early users, coupled with observations of a variety of other users (from the VR experts of the VRAIS ’98 conference to VR novices such as Newt Gingrich and his staff) suggested changes to the system which could increase its effectiveness. Some of these were not implementable due to technology limitations, but the rest were affected.

Some students tried to look at themselves after they had moved through the glass of the interpretive center and out into the gorilla habitat. They were told that when they passed through that barrier that they had “become a gorilla,” and they wanted to examine their gorilla bodies. Since the system had only one tracker to measure head position and orientation, it didn’t have enough information to provide reasonably placed arms and legs. Even if enough trackers had been available, though, providing a gorilla avatar raised new issues: how disorienting would it be to be standing on the demo platform, look down, and see four furry paws on the ground, apparently where your body was supposed to be? Conversely, how confusing would it be to see one’s gorilla avatar behaving as a human.
being, standing on two feet? Unfortunately, lack of sufficient trackers precluded investigating this change.

Sound was a very important part of the system, adding realism and also providing additional cues as to a gorilla’s internal state (the system provided a range of sounds for contented, annoyed, and angry gorillas). In the first prototype system, though, sounds played continuously at a constant volume, no matter where the gorillas were in relation to the student (even if the student was still inside the interpretive center), due to an inability of the SVE toolkit to control the volume of individual sounds separately. Students sometimes found the constant volume confusing, hearing a gorilla rumble and looking around for it since it sounded like it was quite close, even though it was further up the hill. The SVE toolkit was modified so that later versions of the system had sounds that were attenuated using distance, so far away creatures generated lower volume sounds than ones up close. Ideally the system would use spatialized audio so that the student could not only tell from the sound how far away another gorilla was, but also where it was relative to the student. This was implementable on the PC version of SVE, but limitations of the SGI sound library prevented it being implemented there, and in fact made the implementation of volume changes with respect to distance technically difficult. In the interests of cross platform compatibility of the SVE library, spatialized audio was not implemented in the virtual gorilla system.

Some students expressed disappointment that they were not able to actually touch the other gorillas and feel the fur as they were grooming the female. Actually, interactions in the environment were deliberately structured to minimize the need to touch or physically manipulate objects. Since equipment for generalized haptic feedback does not exist, and in fact, providing any haptic feedback at all is still an open research question, all interactions with the virtual gorillas were designed to occur while they were a short distance away from the user. The only interaction allowed with the terrain was to

58
move at a constant height over it. However, gorillas do interact with their environment, playing with sticks or blades of grass, picking up food from the ground, and occasionally touching each other. While it would be easy to have a virtual gorilla interact with a virtual object (for example a stick, or a food item), it would be more of a challenge to provide a way for the student to do the same.

Finally, students seemed to do better when they had a knowledgeable guide to talk them through the first few minutes of interaction with the system. It was expected that they would need a quick introduction to how to look and move around in the virtual environment, so they started out in the virtual interpretive center with someone there to get them used to looking around and moving about inside the building. However, it also proved useful for the guide to remain by their side once they had ventured out into the habitat to answer their questions and talk them through their first interaction with the other gorillas. It was too far outside the students’ experience for them to be able to interpret the sounds and head gestures of the other gorillas without someone asking leading questions to connect what they knew with what they were experiencing, even though they had spent several weeks observing gorilla behavior from outside the habitats.

To address this problem, audio annotations were added that explained features of the habitat design and described what the student was seeing in a manner similar to what the human gorilla experts had provided during the early prototype testing. The initial audio annotations were too verbose, so it was possible for a user to trigger several annotations simultaneously by moving rapidly about the environment. In later revisions, the audio clips were condensed to present only the most important details, and this helped alleviate the problem of multiple clips playing at the same time.

In addition, mood indicators that could be toggled on or off were added above each gorilla to help users discern a gorilla’s mood based on his posture, actions, and vocalizations. Green squares, yellow inverted triangles, red octagons, and white
pensants were used to indicate contentment, annoyance, anger, or submissiveness, respectively.

Quantitative Evaluation

After several revisions to the system based on qualitative evaluations, a more formal quantitative evaluation was undertaken to determine the effectiveness of VR as an educational technology. Two multiple choice tests were devised that tested knowledge of gorillas and their behaviors, based on the educational goals of Zoo Atlanta and the educational objectives of the virtual gorilla system. Each test consisted of 25 multiple choice questions. The first 10 questions showed pictures of a gorilla or gorillas, and asked the user to identify the type of a particular gorilla in the picture. The next six questions had the user play a gorilla vocalization and identify the gorilla mood that it reflected. There was then a question about gorilla habitat design, followed by seven questions about the appropriateness or inappropriateness of various potential gorilla behaviors. The last question concerned the gorilla dominance hierarchy. Each test had the same number of each type of question: four silverback identification questions, three female identification questions, three blackcoal, juvenile, or infant identification questions, four contentment sound questions, one annoyed sound question, one angry sound question, three proximity questions, three gaze questions, and one question about other behaviors, in addition to the habitat and dominance question. The questions for both tests are listed in Appendix A.

Both tests were presented as web forms to enable the students to repeatedly play the sound clips, examine the photos, and redo their answers as they wished. Once the students clicked on the submit answers button, their answers were appended to a file named with their subject number, along with information identifying for which test this set of answers was.

60
An experiment was conducted in which 19 students were given one of the tests and then given the other test directly. Another 21 students were given one of the tests, then allowed to interact with the virtual gorilla environment for up to 20 minutes, and then given the second test. The order of the tests was alternated within the two groups, so 9 students took test A and then test B, 10 students took test B and then test A, 10 students took test A, interacted with the virtual environment, then took test B, and 11 students took test B, interacted with the virtual environment, and then took test A.

Two different tests were used in order to try to assess what a student learned about gorillas and their behaviors. Instead of measuring how many specific answers they had found to the first test questions in the environment when they took the same test a second time, a second test was used that covered the same material in a slightly different fashion. Since the objective was to determine if the virtual environment promoted learning, and not just finding specific facts, the same test was not used for both pretest and posttest. Also, tests were alternated between subjects, so the first subject would take test A first and then test B while the second subject would take test B first and then test A. After completing the experiment, both groups were given a questionnaire to determine how much exposure the subjects had had to virtual environments or gorillas before. This questionnaire is also included in Appendix A.

Experimental subjects were chosen from among students at Oneonta State, who were offered extra credit in various computer science courses for participating. Some students were computer science majors, while others were education majors taking a required computer science elective. The experiment was conducted under the supervision of the Oneonta State Institutional Review Board. Students were asked to read and sign a consent form explaining their rights as subjects, and were given a copy of this form to take with them when leaving. They were also told that they could terminate their participation at any time. In any case, no subject was allowed to remain in the
virtual environment for more than 20 minutes, in order to avoid any potentially detrimental effects that might be caused by long-term exposure to a virtual environment.

The original plan was to have 20 students use the VE and 20 students just take the tests. However, one of the control group subjects who took test A inadvertently didn’t take test B as well, so his results were not included when computing test statistics. In addition, two subjects were inadvertently assigned the same number initially, so an extra student completed the test B, VR, test A phase of the experiment. In the end, nine students took test A and then test B, ten students took test B and then test A, ten students took test A, tried the virtual environment and then took test B, while eleven students took test B, tried the virtual environment, and then took test A.

**Similarity of the Two Tests** The first item to be determined was how similar the two tests were. They were constructed to be as identical as possible without actually asking the exact same questions, but this similarity needed to be determined quantitatively before the results for pretest A could be compared to the results of posttest B, and vice versa.

To determine this, the number of correct answers on test A by those who took test A first was compared with the number of correct answers on test B by those who took test B
first. Figure 4-2 shows the distributions of the pretest scores for the two groups: those that took test A as a pretest and those that took test B as a pretest. Means are illustrated by the solid circles, and outliers by asterisks. As can be seen from the figure, there was an obvious difference in the results of the two pretests, that needed to be investigated statistically.

Table 4.1. Statistical Analysis of the Two Tests

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test A</td>
<td>19</td>
<td>9.63</td>
<td>2.39</td>
</tr>
<tr>
<td>Test B</td>
<td>21</td>
<td>11.71</td>
<td>2.59</td>
</tr>
</tbody>
</table>

Using a two sample t test to compare the means of the two distributions gave the results tabulated in Table 4-1. This showed that the two means had a difference of just over two questions, with a standard deviation of around 2.5 in each case. Comparing the two means using the assumption that they are equal and testing with a 95% confidence interval gave a $P=0.012$, while a 95% confidence interval for the difference of the two means was (-3.68, -0.49). Therefore, with almost a 99% certainty, the two tests were statistically significantly different, and the hypothesis of comparable tests had to be rejected, so it was not enough just to compare the results of pretests with posttests, but they had to be separated out as to which test was taken as pretest as well. In other words, instead of comparing posttest to pretest differences of those who interacted with the virtual environment with those that did not, it was now necessary to compare those who took test A first and interacted with the virtual environment with those that took test A first and didn't interact with the virtual environment. Similarly, those that took test B first and interacted with the virtual environment need to be compared to those that took test B first and didn't use the virtual environment.
One other item of note: no matter the order the tests were administered, none of the subjects answered question 25 correctly on test A, while all of the subjects answered question 25 correctly on test B. This question asked in one instance which was the most dominant gorilla of a list, and in the other which was the least dominant gorilla of the same list. Since subjects always identified the most dominant gorilla correctly and never identified the least dominant gorilla correctly, this question was included when computing statistics for test totals, but was not analyzed by itself.

**Test A Pretest Analysis**: There were 9 subjects who took test A as a pretest and then took test B immediately afterward, and there were 10 subjects who took test A as a pretest, interacted with the virtual environment, and then took test B. To compensate for different incoming levels of gorilla knowledge, an analysis was done of the difference in scores between test A and test B under the assumption that those who knew more about gorillas initially would do better on both tests, so that the difference in scores would help cancel this difference in incoming knowledge. Overall test scores were analyzed, as well as scores for each different type of question. The statistical results are summarized in Tables 4-2 and 4-3, while Table 4-4 provides a brief description of each test task.
<table>
<thead>
<tr>
<th>Test Task</th>
<th>VR?</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silverback ID</td>
<td>N</td>
<td>9</td>
<td>0.66</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.60</td>
<td>1.26</td>
</tr>
<tr>
<td>Female ID</td>
<td>N</td>
<td>9</td>
<td>1.667</td>
<td>0.707</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>1.10</td>
<td>1.29</td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>N</td>
<td>9</td>
<td>-0.33</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>-0.40</td>
<td>0.998</td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>N</td>
<td>9</td>
<td>2.22</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>1.70</td>
<td>1.42</td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>N</td>
<td>9</td>
<td>1.89</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>1.30</td>
<td>1.64</td>
</tr>
<tr>
<td>Contented ID</td>
<td>N</td>
<td>9</td>
<td>0.556</td>
<td>0.822</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.70</td>
<td>0.675</td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>N</td>
<td>9</td>
<td>-0.111</td>
<td>0.601</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.50</td>
<td>0.707</td>
</tr>
<tr>
<td>Angry ID</td>
<td>N</td>
<td>9</td>
<td>0.111</td>
<td>0.333</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.80</td>
<td>0.422</td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>N</td>
<td>9</td>
<td>0.586</td>
<td>0.882</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>2.00</td>
<td>1.05</td>
</tr>
<tr>
<td>Habitat</td>
<td>N</td>
<td>9</td>
<td>0.111</td>
<td>0.501</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.40</td>
<td>0.699</td>
</tr>
<tr>
<td>Proximity</td>
<td>N</td>
<td>9</td>
<td>0.33</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.30</td>
<td>1.16</td>
</tr>
<tr>
<td>Gaze</td>
<td>N</td>
<td>9</td>
<td>0.22</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.10</td>
<td>1.10</td>
</tr>
<tr>
<td>Other Behavior</td>
<td>N</td>
<td>9</td>
<td>-0.222</td>
<td>0.667</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.10</td>
<td>0.588</td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>N</td>
<td>9</td>
<td>3.33</td>
<td>2.12</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.50</td>
<td>1.43</td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td>9</td>
<td>3.89</td>
<td>2.62</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>5.20</td>
<td>3.26</td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>N</td>
<td>9</td>
<td>4.22</td>
<td>2.73</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>5.60</td>
<td>3.06</td>
</tr>
</tbody>
</table>
Table 4-3. Confidence Interval and P Value for Hypothesis $\mu_{\text{ANFR}} = \mu_{\text{AFR}}$ with Pretest A and Posttest B

<table>
<thead>
<tr>
<th>Test Task</th>
<th>P</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silverback ID</td>
<td>0.94</td>
<td>-1.26</td>
<td>1.17</td>
</tr>
<tr>
<td>Female ID</td>
<td>0.25</td>
<td>-0.44</td>
<td>1.58</td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>0.90</td>
<td>-1.02</td>
<td>1.15</td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>0.38</td>
<td>-0.70</td>
<td>1.75</td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>0.38</td>
<td>-0.78</td>
<td>1.96</td>
</tr>
<tr>
<td>Contenst ID</td>
<td>0.70</td>
<td>-0.92</td>
<td>0.63</td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>0.059</td>
<td>-1.25</td>
<td>0.03</td>
</tr>
<tr>
<td>Angry ID</td>
<td>0.0011</td>
<td>-1.06</td>
<td>-0.32</td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>0.0050</td>
<td>-2.39</td>
<td>-0.50</td>
</tr>
<tr>
<td>Habitat</td>
<td>0.35</td>
<td>-0.92</td>
<td>0.34</td>
</tr>
<tr>
<td>Proximity</td>
<td>0.85</td>
<td>-1.13</td>
<td>1.29</td>
</tr>
<tr>
<td>Gaze</td>
<td>0.54</td>
<td>-1.12</td>
<td>1.36</td>
</tr>
<tr>
<td>Other Behavior</td>
<td>0.28</td>
<td>-0.93</td>
<td>0.29</td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>0.85</td>
<td>-1.98</td>
<td>1.85</td>
</tr>
<tr>
<td>Total</td>
<td>0.35</td>
<td>-4.17</td>
<td>1.44</td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>0.32</td>
<td>-4.19</td>
<td>1.44</td>
</tr>
<tr>
<td>Test Task</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silverback ID</td>
<td>Number of silverbacks correctly identified out of 4 images</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female ID</td>
<td>Number of female gorillas correctly identified out of 3 images</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>Number of blackbacks, juveniles, and infants correctly identified out of 3 images</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>Number of silverback and female gorillas correctly identified out of 7 images</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>Number of gorillas correctly identified as to type out of 10 images</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contented ID</td>
<td>Number of contentment vocalizations correctly identified out of 4 sound clips</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>Number of annoyance vocalizations correctly identified out of 1 sound clip</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angry ID</td>
<td>Number of anger vocalizations correctly identified out of 1 sound clip</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>Number of gorilla vocalizations correctly identified out of 6 sound clips</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habitat</td>
<td>Number of habitat questions answered correctly out of 1 question</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximity</td>
<td>Number of questions about acceptable behavior based on proximity answered correctly out of 3 questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaze</td>
<td>Number of questions about acceptable behavior based on gaze answered correctly out of 3 questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Behavior</td>
<td>Number of questions about other acceptable behaviors answered correctly out of 1 question</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acceptable</td>
<td>Number of questions about acceptable behaviors answered correctly out of 7 questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavior Total</td>
<td>Total number of questions answered correctly out of 25 questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>Total number of questions answered correctly out of 22 questions (omitting the 3 questions about identifying other gorillas, since nothing in the environment provided the information needed to answer these)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The statistical analysis revealed a strong trend for VR assisting learning to identify annoyance vocalizations ($P=0.059$), and statistical significance for learning to identify anger vocalizations ($P=0.0011$), and gorilla vocalizations in general ($P=0.0050$), but for all other tasks there was fairly strong to strong evidence to accept the null hypothesis, that there was no improvement on test scores for those who experienced the VR environment between tests and those who did not.

Test B Pretest Analysis: There were 10 subjects who took test B as a pretest and then took test A immediately afterwards, and there were 11 subjects who took test B as a pretest, interacted with the virtual environment, and then took test A. To compensate for different incoming levels of gorilla knowledge, an analysis was done of the difference in scores between test B and test A under the assumption that those who knew more about gorillas initially would do better on both tests, so that the difference in scores would help cancel this difference in incoming knowledge. Overall test scores were analyzed, as well as scores for each different type of question. The statistical results are summarized in Tables 4-5 and 4-6. As before, Table 4-4 contains brief descriptions of the test tasks.
<table>
<thead>
<tr>
<th>Task Task</th>
<th>VR?</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silverback ID</td>
<td>N</td>
<td>10</td>
<td>-0.60</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>-0.45</td>
<td>1.57</td>
</tr>
<tr>
<td>Female ID</td>
<td>N</td>
<td>10</td>
<td>-1.20</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>-1.455</td>
<td>0.522</td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>N</td>
<td>10</td>
<td>0.70</td>
<td>1.57</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.55</td>
<td>1.13</td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>N</td>
<td>10</td>
<td>-1.80</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>-1.91</td>
<td>1.51</td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>N</td>
<td>10</td>
<td>-1.10</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>-1.36</td>
<td>1.75</td>
</tr>
<tr>
<td>Contented ID</td>
<td>N</td>
<td>10</td>
<td>-0.10</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.00</td>
<td>1.18</td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>N</td>
<td>10</td>
<td>0.00</td>
<td>9.00</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.364</td>
<td>0.505</td>
</tr>
<tr>
<td>Angry ID</td>
<td>N</td>
<td>10</td>
<td>0.20</td>
<td>0.632</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.091</td>
<td>0.639</td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>N</td>
<td>10</td>
<td>0.10</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.45</td>
<td>1.44</td>
</tr>
<tr>
<td>Habitat</td>
<td>N</td>
<td>10</td>
<td>0.10</td>
<td>0.738</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.091</td>
<td>0.831</td>
</tr>
<tr>
<td>Proximity</td>
<td>N</td>
<td>10</td>
<td>-0.20</td>
<td>0.632</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.64</td>
<td>1.12</td>
</tr>
<tr>
<td>Gaze</td>
<td>N</td>
<td>10</td>
<td>0.00</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>-0.18</td>
<td>1.08</td>
</tr>
<tr>
<td>Other Behavior</td>
<td>N</td>
<td>10</td>
<td>0.20</td>
<td>0.632</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.91</td>
<td>0.539</td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>N</td>
<td>10</td>
<td>0.00</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.55</td>
<td>1.75</td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td>10</td>
<td>-1.90</td>
<td>2.18</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>-1.27</td>
<td>3.38</td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>N</td>
<td>10</td>
<td>-2.60</td>
<td>1.84</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>-1.82</td>
<td>2.89</td>
</tr>
<tr>
<td>Test Task</td>
<td>P</td>
<td>95% Confidence Interval for $\mu_{\text{pre}} - \mu_{\text{post}}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>----</td>
<td>---------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silverback ID</td>
<td>0.62</td>
<td>-1.45</td>
<td>1.16</td>
<td></td>
</tr>
<tr>
<td>Female ID</td>
<td>0.50</td>
<td>-0.53</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>0.80</td>
<td>-1.12</td>
<td>1.43</td>
<td></td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>0.87</td>
<td>-1.23</td>
<td>1.44</td>
<td></td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>0.77</td>
<td>-1.60</td>
<td>2.12</td>
<td></td>
</tr>
<tr>
<td>Contented ID</td>
<td>0.84</td>
<td>-1.10</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Angry ID</td>
<td>0.68</td>
<td>-0.43</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>0.53</td>
<td>-1.52</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Habitat</td>
<td>0.98</td>
<td>-0.71</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>Proximity</td>
<td>0.049</td>
<td>-1.67</td>
<td>-0.09</td>
<td></td>
</tr>
<tr>
<td>Gaze</td>
<td>0.75</td>
<td>-0.99</td>
<td>1.35</td>
<td></td>
</tr>
<tr>
<td>Other Behavior</td>
<td>0.66</td>
<td>-0.43</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>0.45</td>
<td>-2.03</td>
<td>3.94</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.62</td>
<td>-3.22</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>0.47</td>
<td>-2.99</td>
<td>1.43</td>
<td></td>
</tr>
</tbody>
</table>

* Note that NonVR subjects scored the same on pretest and posttest for this question.
One item to be noted is that in the group that did not experiment with the virtual environment, none of them correctly identified the annoyance vocalization either in the pretest or in the posttest. This made it impossible to compute a P value, but since more of the VR group did identify the annoyance vocalization after interacting with the virtual environment than before, there is a significant difference between the two groups for this task.

The statistical analysis revealed statistical significance for VR assisting learning to identify socially acceptable behavior when near other gorillas (P<0.049), and as noted above, apparently for identifying annoyance vocalizations, but for all other tasks there was fairly strong to strong evidence to accept the null hypothesis, that there was no improvement on test scores for those who experienced the VR environment between tests and those who did not.

Posttest A Analysis. One possibility for the lack of significant results was that the differences between Test A and Test B were so large that they swamped any differences between the VR and non-VR groups. To see if this might have been the case, the posttest results were looked at by themselves, instead of examining the differences between posttest and pretest results. While this would eliminate the correction for differing initial amounts of knowledge about gorillas, it would also remove the disparity between tests as a factor in the statistical analysis. This analysis was undertaken next. Ten subjects took the test A posttest without any interactions with the virtual gorilla environment, while eleven subjects took the test A posttest after exploring the virtual environment. The statistical results of this analysis are contained in Table 4-7 and Table 4-8.
<table>
<thead>
<tr>
<th>Test Task</th>
<th>VR?</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silverback ID</td>
<td>N</td>
<td>10</td>
<td>2.10</td>
<td>0.876</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>1.55</td>
<td>1.21</td>
</tr>
<tr>
<td>Female ID</td>
<td>N</td>
<td>10</td>
<td>0.80</td>
<td>0.789</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.727</td>
<td>0.647</td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>N</td>
<td>10</td>
<td>2.00</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>1.36</td>
<td>1.12</td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>N</td>
<td>10</td>
<td>2.90</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>2.27</td>
<td>1.62</td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>N</td>
<td>10</td>
<td>4.90</td>
<td>1.85</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>3.64</td>
<td>1.43</td>
</tr>
<tr>
<td>Contented ID</td>
<td>N</td>
<td>10</td>
<td>0.60</td>
<td>0.843</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.727</td>
<td>0.905</td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>N</td>
<td>10</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.455</td>
<td>0.522</td>
</tr>
<tr>
<td>Angry ID</td>
<td>N</td>
<td>10</td>
<td>0.30</td>
<td>0.483</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.273</td>
<td>0.467</td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>N</td>
<td>10</td>
<td>0.90</td>
<td>0.876</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>1.45</td>
<td>1.13</td>
</tr>
<tr>
<td>Habitat</td>
<td>N</td>
<td>10</td>
<td>0.40</td>
<td>0.518</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.455</td>
<td>0.522</td>
</tr>
<tr>
<td>Proximity</td>
<td>N</td>
<td>10</td>
<td>1.30</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>2.364</td>
<td>0.505</td>
</tr>
<tr>
<td>Gaze</td>
<td>N</td>
<td>10</td>
<td>1.50</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>1.545</td>
<td>0.820</td>
</tr>
<tr>
<td>Other Behavior</td>
<td>N</td>
<td>10</td>
<td>0.90</td>
<td>0.316</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>0.909</td>
<td>0.302</td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>N</td>
<td>10</td>
<td>3.70</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>4.82</td>
<td>1.08</td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td>10</td>
<td>9.90</td>
<td>3.14</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>10.36</td>
<td>2.54</td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>N</td>
<td>10</td>
<td>7.90</td>
<td>3.03</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>11</td>
<td>9.00</td>
<td>2.19</td>
</tr>
</tbody>
</table>
Table 4-8. Confidence Interval and P Value for Hypothesis $\mu_{\text{post}} = \mu_{\text{pre}}$, with Posttest A

<table>
<thead>
<tr>
<th>Test Task</th>
<th>P</th>
<th>95% Confidence Interval for $\mu_{\text{post}} - \mu_{\text{pre}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silverback ID</td>
<td>0.24</td>
<td>Min -0.41, Max 1.52</td>
</tr>
<tr>
<td>Female ID</td>
<td>0.82</td>
<td>Min -0.60, Max 0.74</td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>0.24</td>
<td>Min -0.45, Max 1.73</td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>0.32</td>
<td>Min -0.67, Max 1.92</td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>0.10</td>
<td>Min -0.28, Max 2.81</td>
</tr>
<tr>
<td>Contented ID</td>
<td>0.74</td>
<td>Min -0.93, Max 0.67</td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Angry ID</td>
<td>0.90</td>
<td>Min -0.41, Max 0.46</td>
</tr>
<tr>
<td>Vocalization+ ID</td>
<td>0.22</td>
<td>Min -1.48, Max 0.37</td>
</tr>
<tr>
<td>Habitual</td>
<td>0.81</td>
<td>Min -0.53, Max 0.42</td>
</tr>
<tr>
<td>Proximity</td>
<td>0.0075</td>
<td>Min -1.79, Max -0.34</td>
</tr>
<tr>
<td>Gaze</td>
<td>0.93</td>
<td>Min -1.15, Max 1.06</td>
</tr>
<tr>
<td>Other Behavior</td>
<td>0.95</td>
<td>Min -0.22, Max 0.275</td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>0.14</td>
<td>Min -2.56, Max 0.42</td>
</tr>
<tr>
<td>Total</td>
<td>0.72</td>
<td>Min -3.11, Max 2.18</td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>0.36</td>
<td>Min -3.57, Max 1.37</td>
</tr>
</tbody>
</table>

* Note that NonVR subjects all missed this question.

Considering just the results of posttest A, statistical significance supports the hypothesis that students learned to identify gorilla annoyance vocalizations, and what was acceptable behavior when in close proximity to another gorilla. These are the same two areas that showed significance when analyzing the score differences between posttest A and pretest B.

**Posttest B Analysis** Nine subjects took the test B posttest without any interactions with the virtual gorilla environment, while ten subjects took the test B posttest after exploring the virtual environment. The statistical results of this analysis are contained in Table 4-9 and Table 4-10.
<table>
<thead>
<tr>
<th>Test Task</th>
<th>VR?</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silverback ID</td>
<td>N</td>
<td>9</td>
<td>2.556</td>
<td>0.726</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>2.60</td>
<td>0.643</td>
</tr>
<tr>
<td>Female ID</td>
<td>N</td>
<td>9</td>
<td>2.222</td>
<td>0.667</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>2.20</td>
<td>0.632</td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>N</td>
<td>9</td>
<td>1.444</td>
<td>0.527</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.90</td>
<td>0.568</td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>N</td>
<td>9</td>
<td>4.78</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>4.80</td>
<td>1.14</td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>N</td>
<td>9</td>
<td>6.222</td>
<td>0.972</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>5.70</td>
<td>1.16</td>
</tr>
<tr>
<td>Contented ID</td>
<td>N</td>
<td>9</td>
<td>1.44</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>1.10</td>
<td>0.738</td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>N</td>
<td>9</td>
<td>0.11</td>
<td>0.333</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.60</td>
<td>0.516</td>
</tr>
<tr>
<td>Angry ID</td>
<td>N</td>
<td>9</td>
<td>0.11</td>
<td>0.333</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.90</td>
<td>0.316</td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>N</td>
<td>9</td>
<td>1.67</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>2.60</td>
<td>0.966</td>
</tr>
<tr>
<td>Habitat</td>
<td>N</td>
<td>9</td>
<td>0.444</td>
<td>0.527</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.80</td>
<td>0.422</td>
</tr>
<tr>
<td>Proximity</td>
<td>N</td>
<td>9</td>
<td>1.778</td>
<td>0.833</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>2.00</td>
<td>0.943</td>
</tr>
<tr>
<td>Gaze</td>
<td>N</td>
<td>9</td>
<td>1.778</td>
<td>0.833</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>1.80</td>
<td>0.632</td>
</tr>
<tr>
<td>Other Behavior</td>
<td>N</td>
<td>9</td>
<td>0.667</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>0.90</td>
<td>0.316</td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>N</td>
<td>9</td>
<td>4.222</td>
<td>0.972</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>4.70</td>
<td>1.25</td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td>9</td>
<td>13.56</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>14.80</td>
<td>2.62</td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>N</td>
<td>9</td>
<td>12.11</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>10</td>
<td>13.90</td>
<td>2.64</td>
</tr>
</tbody>
</table>
Table 4-10. Confidence Interval and P Value for Hypothesis $\mu_{\text{pre}} = \mu_{\text{post}}$
with Posttest B

<table>
<thead>
<tr>
<th>Test Task</th>
<th>P</th>
<th>95% Confidence Interval for $\mu_{\text{pre}} - \mu_{\text{post}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silverback ID</td>
<td>0.90</td>
<td>Min: -0.81, Max: 0.72</td>
</tr>
<tr>
<td>Female ID</td>
<td>0.94</td>
<td>Min: -0.61, Max: 0.66</td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>0.046</td>
<td>Min: 0.01, Max: 1.08</td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>0.97</td>
<td>Min: -1.11, Max: 1.06</td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>0.30</td>
<td>Min: -0.51, Max: 1.56</td>
</tr>
<tr>
<td>Contented ID</td>
<td>0.45</td>
<td>Min: -0.61, Max: 1.39</td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>0.026</td>
<td>Min: -0.97, Max: -0.07</td>
</tr>
<tr>
<td>Angry ID</td>
<td>0.0001</td>
<td>Min: -1.11, Max: -0.47</td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>0.072</td>
<td>Min: -1.96, Max: 0.09</td>
</tr>
<tr>
<td>Habitat</td>
<td>0.13</td>
<td>Min: 3.63, Max: 0.11</td>
</tr>
<tr>
<td>Proximity</td>
<td>0.59</td>
<td>Min: -1.09, Max: 0.64</td>
</tr>
<tr>
<td>Gaze</td>
<td>0.95</td>
<td>Min: -0.76, Max: 0.71</td>
</tr>
<tr>
<td>Other Behavior</td>
<td>0.25</td>
<td>Min: -0.65, Max: 0.19</td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>0.36</td>
<td>Min: -1.56, Max: 0.61</td>
</tr>
<tr>
<td>Total</td>
<td>0.20</td>
<td>Min: -3.24, Max: 0.75</td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>0.085</td>
<td>Min: -3.86, Max: 0.28</td>
</tr>
</tbody>
</table>

Analyzing just the results of posttest B, at the 0.05 significance level there was a
significant difference after exploring the virtual environment in identifying other gorillas
(blackbacks, juveniles, and infants) (P=0.046), in identifying gorilla annoyance
vocalizations (P=0.026), and in identifying gorilla anger vocalizations (P=0.0001). While
the results for general vocalization identification (P=0.072) and all test answers with the
exception of other gorilla identification (P=0.085) showed a trend, the results are not
statistically significant.

These results differ somewhat from the results obtained when comparing posttest
B minus pretest A scores, since in that case, the anger and general vocalization
identification results showed significance, while the other gorilla identification results were
definitely not significant, and the results of the annoyance identification were interesting
but not conclusively significant.

75
Examining the posttest results and comparing them with the analysis of the posttest minus pretest results there are strong similarities, but it becomes more apparent that the test mean disparities are masking some of the differences between VR and nonVR users.

Test A Results Comparisons. One other possible method of analyzing the results is to just consider one of the tests, and to compare the scores of those who took it as a pretest with the scores of those who took it as a posttest. Assuming the two sample spaces were the same, then this comparison would show if learning occurred when using the virtual environment without needing to compensate in the computation for the differences between the two tests. This analysis was undertaken next, comparing the population of 19 people who took test A as a pretest with the 11 people who took test A as a posttest after using the virtual environment. If the test scores were higher after using the virtual environment, that would be another indicator of the efficacy of virtual reality as an educational tool.

The results of this analysis are summarized in Tables 4-11 and 4-12.
<table>
<thead>
<tr>
<th>Test Task</th>
<th>Pre or Post?</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silverback ID</td>
<td>Pre</td>
<td>19</td>
<td>2.00</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>1.55</td>
<td>1.21</td>
</tr>
<tr>
<td>Female ID</td>
<td>Pre</td>
<td>19</td>
<td>0.842</td>
<td>0.834</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>0.727</td>
<td>0.647</td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>Pre</td>
<td>19</td>
<td>1.53</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>1.26</td>
<td>1.12</td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>Pre</td>
<td>19</td>
<td>2.84</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>2.27</td>
<td>1.52</td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>Pre</td>
<td>19</td>
<td>4.37</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>3.64</td>
<td>1.43</td>
</tr>
<tr>
<td>Contented ID</td>
<td>Pre</td>
<td>19</td>
<td>0.632</td>
<td>0.684</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>0.727</td>
<td>0.905</td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>Pre</td>
<td>19</td>
<td>0.158</td>
<td>0.375</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>0.455</td>
<td>0.522</td>
</tr>
<tr>
<td>Angry ID</td>
<td>Pre</td>
<td>19</td>
<td>0.053</td>
<td>0.229</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>0.273</td>
<td>0.467</td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>Pre</td>
<td>19</td>
<td>0.842</td>
<td>0.765</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>1.45</td>
<td>1.13</td>
</tr>
<tr>
<td>Habitat</td>
<td>Pre</td>
<td>19</td>
<td>0.388</td>
<td>0.498</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>0.455</td>
<td>0.522</td>
</tr>
<tr>
<td>Proximity</td>
<td>Pre</td>
<td>19</td>
<td>1.579</td>
<td>0.838</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>2.364</td>
<td>0.505</td>
</tr>
<tr>
<td>Gaze</td>
<td>Pre</td>
<td>19</td>
<td>1.632</td>
<td>0.631</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>1.545</td>
<td>0.820</td>
</tr>
<tr>
<td>Other Behavior</td>
<td>Pre</td>
<td>19</td>
<td>0.842</td>
<td>0.375</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>0.909</td>
<td>0.302</td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>Pre</td>
<td>19</td>
<td>4.05</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>4.02</td>
<td>1.08</td>
</tr>
<tr>
<td>Total</td>
<td>Pre</td>
<td>19</td>
<td>9.63</td>
<td>2.39</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>10.36</td>
<td>2.54</td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>Pre</td>
<td>19</td>
<td>8.11</td>
<td>2.38</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>11</td>
<td>9.00</td>
<td>2.18</td>
</tr>
</tbody>
</table>
Comparing scores on test A taken as a pretest with scores on test B taken as a posttest after exploring the virtual gorilla habitat, the only result of significance is learning about acceptable behaviors when passing close to another gorilla (P=0.0035).

**Test B Results Comparisons**  In a similar fashion, the scores of the 21 people who took test B as a pretest can be compared to the scores of the 10 people who took test B as a posttest. These results are summarized in Table 4-13 and 4-14.
<table>
<thead>
<tr>
<th>Test Task</th>
<th>Pre or Post?</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silverback ID</td>
<td>Pre</td>
<td>21</td>
<td>2.333</td>
<td>0.966</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>2.600</td>
<td>0.843</td>
</tr>
<tr>
<td>Female ID</td>
<td>Pre</td>
<td>21</td>
<td>2.095</td>
<td>0.700</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>2.200</td>
<td>0.632</td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>Pre</td>
<td>21</td>
<td>1.048</td>
<td>0.498</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>0.90</td>
<td>0.568</td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>Pre</td>
<td>21</td>
<td>4.43</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>4.80</td>
<td>1.14</td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>Pre</td>
<td>21</td>
<td>5.48</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>5.70</td>
<td>1.16</td>
</tr>
<tr>
<td>Contented ID</td>
<td>Pre</td>
<td>21</td>
<td>0.714</td>
<td>0.644</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>1.10</td>
<td>0.738</td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>Pre</td>
<td>21</td>
<td>0.048</td>
<td>0.218</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>0.600</td>
<td>0.516</td>
</tr>
<tr>
<td>Angry ID</td>
<td>Pre</td>
<td>21</td>
<td>0.143</td>
<td>0.369</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>0.900</td>
<td>0.316</td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>Pre</td>
<td>21</td>
<td>0.965</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>2.80</td>
<td>0.966</td>
</tr>
<tr>
<td>Habitat</td>
<td>Pre</td>
<td>21</td>
<td>0.333</td>
<td>0.483</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>0.80</td>
<td>0.422</td>
</tr>
<tr>
<td>Proximity</td>
<td>Pre</td>
<td>21</td>
<td>1.619</td>
<td>0.805</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>2.00</td>
<td>0.943</td>
</tr>
<tr>
<td>Gaze</td>
<td>Pre</td>
<td>21</td>
<td>1.619</td>
<td>0.805</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>1.80</td>
<td>0.632</td>
</tr>
<tr>
<td>Other Behavior</td>
<td>Pre</td>
<td>21</td>
<td>0.762</td>
<td>0.436</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>0.90</td>
<td>0.316</td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>Pre</td>
<td>21</td>
<td>4.30</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>4.70</td>
<td>1.25</td>
</tr>
<tr>
<td>Total</td>
<td>Pre</td>
<td>21</td>
<td>11.71</td>
<td>2.59</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>14.80</td>
<td>2.82</td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>Pre</td>
<td>21</td>
<td>10.67</td>
<td>2.54</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>10</td>
<td>13.90</td>
<td>2.84</td>
</tr>
<tr>
<td>Test Task</td>
<td>P</td>
<td>$\mu_{\text{result}}$ - $\mu_{\text{baseline}}$</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------</td>
<td>----------------------------------------------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Silverback ID</td>
<td>0.44</td>
<td>-0.99</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>Female ID</td>
<td>0.68</td>
<td>-0.63</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>Other Gorilla ID</td>
<td>0.49</td>
<td>-0.30</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Silverback + Female ID</td>
<td>0.39</td>
<td>-1.27</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>Gorilla ID</td>
<td>0.62</td>
<td>-1.17</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>Contented ID</td>
<td>0.18</td>
<td>-0.57</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Annoyed ID</td>
<td>0.0088</td>
<td>-0.931</td>
<td>-0.17</td>
<td></td>
</tr>
<tr>
<td>Angry ID</td>
<td>0.0000</td>
<td>-1.022</td>
<td>-0.49</td>
<td></td>
</tr>
<tr>
<td>Vocalizations ID</td>
<td>0.0003</td>
<td>-2.47</td>
<td>-3.92</td>
<td></td>
</tr>
<tr>
<td>Habitat</td>
<td>0.012</td>
<td>-0.82</td>
<td>-0.11</td>
<td></td>
</tr>
<tr>
<td>Proximity</td>
<td>0.29</td>
<td>-1.12</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>Gaze</td>
<td>0.50</td>
<td>-0.73</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Other Behavior</td>
<td>0.33</td>
<td>-0.424</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Acceptable Behavior</td>
<td>0.18</td>
<td>-1.75</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.0068</td>
<td>-5.20</td>
<td>-0.97</td>
<td></td>
</tr>
<tr>
<td>Adjusted Total</td>
<td>0.0050</td>
<td>-5.35</td>
<td>-1.12</td>
<td></td>
</tr>
</tbody>
</table>

Examining these results reveals a significant change between groups in identifying annoyed vocalizations (P=0.0088), angry vocalizations (P=0.0000), and general gorilla vocalizations (P=0.0003). In addition, there was a significant difference in information known about gorilla habitats (P=0.012), as well as overall in the number of questions answered correctly (P=0.0068) and the number of questions answered correctly, excluding questions about identifying blackbacks, juveniles, and infants (P=0.005).

So what did quantitative testing say about the virtual gorilla environment, and specifically about the use of virtual environments for concept acquisition in education? This will be examined in the next chapter.
CHAPTER 5
DISCUSSION AND CONCLUSIONS

Discussion

As the analysis in the previous chapter showed, users of the virtual gorilla system learned some facts about gorillas and their social interactions, but other concepts that the system was expected to teach effectively were not conveyed to the students. The student testers were excited about getting to use a virtual environment, and many of them stayed in the environment and explored as long as they were allowed to do so. Clearly they found the environment engaging. Yet many of the concepts the system was designed to teach evidently did not sink in as the students interacted with the system.

The most successful part of the system was the portion focusing on gorilla vocalizations. Statistically significant improvements in vocalization recognition occurred after the students interacted with the virtual environment. Conversely, students did not significantly improve their recognition of different gorilla types, even after interacting with them for upwards of twenty minutes. This contrast might contain a clue to why students learned some things but not others when exploring the virtual environment.

The gorilla vocalizations used in the environment were vocalizations of the actual gorillas at Zoo Atlanta. Although vocalizations of several different gorillas were spliced together to generate a generic contentment vocalization, all vocalizations were generated by real gorillas. (It is interesting to note that many sound effects houses do not sell actual gorilla vocalizations, but instead sell tapes of humans making gorilla sounds as “authentic” gorilla vocalizations.) Although initially some students thought the contentment vocalizations were threatening, they became accustomed to them as they continued to explore the environment.

81
In contrast, while the gorilla models were based on the best available measurements at the time, they were crude approximations to the appearance of an actual gorilla. Because of the need to keep the polygon count down, the models were rather blocky and simplified. In addition, the models were constructed using an old version of Wavefront modeling software, which supported only rudimentary texture mapping. Although much effort was invested in applying some kind of fur texture to the models, the results were never good enough to be used in the system.

One possibility, then is that the verisimilitude of the model is an important factor in determining whether learning occurs or not. It could be that the better the model is, the more learning that occurs, or it could be that there is a threshold beyond which learning will occur and below which it will not. This is something that could use further investigation, and will be discussed in more detail below in the future directions section.

Some students didn’t achieve the full range of interactions, even though they stayed in the virtual environment for ten or fifteen minutes. They would approach a gorilla, and when it would start to stand up and interact with them, they would run away. They did not experience the full gamut of gorilla vocalizations and emotions. This could be an age-related difference, since informal observations of middle school users showed them to be quite willing to explore and try things out, while older students were more reluctant to “do something wrong”, and so were less likely to test the boundaries of the system. Clearly some way to encourage students to experience the full range of interactions, and to explore the environment more completely would be useful.

On the other hand, some of the younger students would challenge the silverback for dominance repeatedly, even though they kept getting put in “time out” because “...it’s fun!” Other students were not totally at ease with the concept of just exploring since it was too open ended. They kept asking “...how do you win?” and wanting more explicit
goals. In the initial tests, a gorilla expert from Zoo Atlanta talked the students through their experiences, encouraging the reluctant explorers while reining in the more exuberant ones. The verbal annotations were added to the original system to try to capture that expertise and guidance, but clearly this is a fruitful area for further investigation.

Observing the students interact with an early version of the system that contained audio annotations showed that some students initially would scurry around the environment, giving everything a cursory examination before settling in to explore each area in depth. Since the audio annotations were triggered either by proximity or gaze, the result would be a long queue of audio clips waiting to be played. Most of the clips would end up being played without the corresponding context, causing the student to ignore them. Also, several of the audio annotations were set to play only once when first triggered, since it would be distracting to keep repeating them. This resulted in the students missing many of the more useful explanations of the environment and the interactions, and reduced the learning opportunities.

To combat this problem, the audio annotations were edited for brevity and conciseness, eliminating complete sentences when a terse phrase would do as well. The resulting annotations are listed in appendix B. This, and shrinking the areas that triggered annotations, helped eliminate the problem of queued audio annotations, but resulted in comments that at times were almost too terse. Building a more intelligent audio annotation system would be another area for further research.

Conclusions

The results of this study were similar to those of the investigations reviewed in the second chapter. While students did succeed in learning some concepts about gorillas, their behaviors, and their social interactions, other concepts that were designed
into the system were not conveyed as successfully. It still remains to be discovered the
guiding principles that determine what will and will not be successful in an educational
virtual environment.

In addition, the process of constructing a complete virtual environment is tedious
and quite time consuming. Better tools, such as modeling environments or VR toolkits,
are becoming available which help ease this task, but it is still a significant amount of
work to construct an educational VE to convey a single set of related concepts. For
virtual reality to impact the classroom, there need to be many educational units
constructed so that the cost of the equipment is amortized over many different learning
objectives. Given that it is still unclear as to why one feature of an educational VE is
successful while another seemingly similar one is not, people are not ready to commit to
the effort required to generate enough educational content to make a VR system a
useful classroom adjunct. This is an area that desperately needs immediate
investigation in order to advance the field. Until then, educational VEs will remain in the
labs and museums of the world.

Along the same lines, while all students enjoyed their time in the virtual gorilla
environment, some learned more than others. The novelty of the experience kept the
students focused on their explorations and interactions, but that novelty at times
distracted students from the material to be learned. Eventually the novelty will wear off,
and this potentially will improve the learning experience since the students will then be
able to focus more on their explorations and interactions with the environment, and less
on the novelty of using an HMD or tracker. Until that time, though, educational VEs
need to make sure the students are presented with the material to be mastered no
matter what path they choose to take when exploring the VE.
Finally, one last surprise from this experiment was the difficulty in successfully testing the learning that occurred when students explored the virtual gorilla environment. Despite the care taken in the construction of the two tests to insure that each was as similar as possible to the other without asking the identical questions, there was a strong statistical difference in the results of the two tests. As the EVL found when they installed an ImmersaDesk in an elementary school\textsuperscript{23}, constructing an educational virtual environment is just the first challenge. A further challenge is to design a way to evaluate its effectiveness accurately.

**Future Directions**

The long term goal of researchers studying educational virtual reality is to discern the larger patterns, and tease out the rules of what is effective and what is ineffective when constructing educational virtual environments. With sufficient exploration of the space of possible educational virtual environments, it is possible that characteristics in common among all successful ones will be discerned. The hope is that these principles exist and may be explicitly defined, so that future educational VEs can be evaluated for their effectiveness before they are constructed, and their designs can be guided by these principles in order that the VEs may be as effective as possible. Many further studies are needed and more systems constructed until these trends become apparent.

With that in mind, the following items are proposed as future work to provide a small first step towards this ultimate goal.

- **Gorilla Identification** was one area where the virtual gorilla environment failed. One possible explanation was that the virtual gorillas did not look enough like real ones for the students to be able to map from the VE to actual gorilla photographs. Since computer graphics hardware and software have continued to improve at an exponential pace, this hypothesis is finally able to be tested further. The latest generation of
graphics cards support a much higher polygon rendering rate, and also add support for pixel and vertex shaders. In addition, the latest generation of modeling software packages allow textures to be "painted" on to models in nearly any way imaginable, instead of restricting the user to simple techniques such as spherical or cylindrical mapping. It should now be possible to build and use much more detailed gorilla models without seriously impacting the system throughput. An interesting follow-on to this work would be to upgrade the gorilla models and redo the experiment to see if student performance on gorilla identification would significantly improve.

Since some have speculated that photorealism can actually detract from learning, such an environment should be compared to the original one but also a third one in which the gorillas are merely caricatures of real gorillas, overly emphasizing the key features that distinguish the various types. Comparing the three systems to see which would best train users to identify the different types of gorillas could shed some illumination on the question of how much "realism" is necessary in an educational VR system in order for learning to take place.

Another important spectrum of values that needs exploring is the gamut from free exploration to totally scripted interactions. Allowing users free rein lets the user focus on areas of interest and topics that are at the frontiers of their knowledge. This should allow each user to customize his experience to best fit his current knowledge. On the other hand, allowing users complete control over their environmental interactions greatly increases the likelihood that users will not completely experience all the features of the VE, and thus will miss out on some of the learning opportunities contained within. Hopefully there is a balance between the two extremes which optimizes the learning while allowing users to personalize the experience to their current knowledge. It would
be even more wonderful if that balance occurred at approximately the same place for all virtual environments, although that is probably too much to expect.

To evaluate this further, the virtual gorilla environment could be modified so that user exploration occurred along a prescribed path, devised to expose the student to all learning opportunities in the environment. A second modification could be made to allow the users to explore freely but over time to make increasingly strong suggestions to the student that he investigate parts of the environment he has overlooked so far. Comparing the learning that occurs in these three systems might help suggest the appropriate level of free will that should be allowed in an educational VE.

From their exposure to video games, students have certain expectations of any computer-generated environment. They expect active rewards when they succeed at some task, even one as benign as interacting properly with other gorillas in the environment. After maintaining peace within the family group for a minute or two, several students asked “...is this all there is to it? How do I win?” Apparently the passive reward of maintaining a peaceful habitat and not getting sent to time out was not enough of a reward for students used to games in which they are awarded rankings, and the opportunity to make the top ten scoreboard. Also, students were inquisitive as to whether or not they could assume other gorilla personas.

Students entered the virtual gorilla environment as juveniles since these are the gorillas in the gorilla family group most like the students; that is, they are entering new relationships with their group as they are transitioning from infant to adult, and are testing the waters to determine the rules governing appropriate behavior. By being powerless, students were forced to conform to proper etiquette or suffer the consequences imposed by the other adults. A deliberate design decision was made not
to let students be the reigning silverback since as such they could impose distinctly non-
gorilla behavior standards on the rest of the group.

One possible solution to the passive reward structure of the environment and the
unfulfilled desire to assume other gorilla roles would be to modify the environment so
that the student would start as a juvenile. As a reward for exhibiting proper gorilla
etiquette while exploring, the student could gain experience points. When enough of
these points were accumulated, the student could be given the option of transitioning to
a slightly more dominant gorilla. The ultimate reward for continuing to demonstrate
proper gorilla social interactions would be a chance to assume the role of the senior
silverback and to rule the rest of the group. This would provide additional
encouragement for proper behavior, would allow the students to become other gorillas
besides a juvenile as they demonstrated a deeper understanding of gorilla social
interactions, and would provide a stronger reward for proper gorilla behavior than just
not getting sent to “time out.”

In addition to the direct extensions to the environment to enhance learning
described above, there are some technical improvements that could be made to the
environment that would indirectly improve the student experience as well. The behavior
controller as implemented works for antagonistic social interactions aimed at reinforcing
the dominance hierarchy. However, gorillas also have many affiliative social interactions
as well. These include play, grooming, and just hanging out together. A true test of the
behavior controller would be to add affiliative behaviors as well as antagonistic ones, to
see if the suggested controller is scalable to handle more as well as more types of
social interactions. In addition, this would make the environment more diverse and
interesting to the user, giving them a wider range of experiences to explore, and
providing a wider range of information about different types of gorilla social interactions.
Finally, anecdotal evidence suggests that having an actual gorilla expert talk
students through their explorations of the virtual gorilla world proved more helpful than
the current system of audio annotations which tried to capture the same experience.
Clearly a better system of annotations that more completely captures the interactions
with the gorilla expert are needed for the virtual gorilla system to reach its full potential
as an educational environment. This means that more intelligence needs to be built into
the annotation system to sense when the user is just looking around instead of being
confused, and to tailor the level of explanation accordingly. While it will not be possible
any time soon to build an agent that can seamlessly replace the gorilla expert, it should
be possible to greatly improve the current system. The first step would be to video tape
user interactions with a gorilla expert and user interactions with the current system to
look for similarities and differences, to determine when the gorilla expert was more
helpful than the computerized system, and to try to extract common patterns for help
triggers that the gorilla expert was picking up that might be detectable by the
computerized system. A lot of interesting work remains to be done just on this problem
alone.

In summary, as is often the case, this study answers its central question of the
utility of a virtual environment for teaching about gorilla behaviors with a resounding
maybe. It also raises as many questions as it answers, questions whose answers will
help expand the frontiers of our understanding about the educational utility of virtual
environments in general, and for teaching about animal behaviors in particular. Many
people believe that it makes sense that virtual environments could prove to be great
educational tools. It remains for those of us interested in these issues to determine
under what conditions this is true, and how to maximize the learning that takes place in
any educational VE.
1. http://www.i-glasses.com/ where the company has been reborn
2. http://www.iisvr.com/products VR vFx3d.html where the VFX3D is listed as a discontinued product
3. http://www.sony.co.jp/en/SonyInfo/News/Press/199809/98-101/ which is the initial press release from Sony for the GLASSTRON, and http://www.sony-style.com.my/tvapp/glastro.png which was the only Sony site listing any version of the GLASSTRON, and lists the GLASSTRON as sold out
4. http://www.virtuality.com/toc/kgtslib/index.html where Cybermind (http://www.cybermind.co.uk/master.html) seems to have acquired the location based entertainment product line of Virtuality
5. Kevin J. Anderson and Doug Beason, Virtual Destruction, 1996
13. http://www.soe.ecu.edu/vrl/vphib.html by Veronica Pantelidis, professor and co-director of the Virtual Reality and Education Laboratory, School of Education, East Carolina University, in Greenville, NC.
14. personal communication from an anonymous referee, VR 2002 conference
17. Howard Rheingold, Virtual Reality, page 45.
18. John T. Bruer, Schools for Thought, 1997,

91


APPENDIX A
SURVEY INSTRUMENTS

Test A

1. The gorilla pictured standing above is a:
   o Silverback
   o Blackback
   o Female
   o Juvenile
   o Infant

2. The gorilla pictured on the right is a:
   o Silverback
   o Blackback
   o Female
   o Juvenile
   o Infant
3. The gorilla pictured sitting above is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant

4. The gorilla pictured above with another gorilla on its back is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant
5. The gorilla pictured sitting above is a:
   o Silverback
   o Blackback
   o Female
   o Juvenile
   o Infant

6. The gorilla pictured walking above is a:
   o Silverback
   o Blackback
   o Female
   o Juvenile
   o Infant

95
7. The gorilla pictured foraging above is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant

8. The gorilla pictured above riding on the back of the other gorilla is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant
9. The gorilla pictured above eating broccoli is a:
   o Silverback
   o Blackback
   o Female
   o Juvenile
   o Infant

10. The gorilla pictured standing above is a:
    o Silverback
    o Blackback
    o Female
    o Juvenile
    o Infant
For the next series of questions, click on the speaker icon to hear the sound for the question below the icon.

11. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear

12. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear

13. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear

14. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear
15. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear

16. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear

17. The moats around the habitat provide:
   - A place for gorillas to practice their climbing skills
   - Running water for the gorillas to play in and drink
   - A simulation of the dry river beds in Africa
   - A barrier that doesn't obstruct one's view
   - A shady place to get out of the sun

For each of the following activities or actions, is the action generally appropriate or inappropriate gorilla behavior for unrelated gorillas?

18. A female walking within one foot of a juvenile
   - Socially appropriate gorilla behavior
   - Socially inappropriate gorilla behavior

19. A female staring at a silverback who is looking back at her from 15 feet away
   - Socially appropriate gorilla behavior
   - Socially inappropriate gorilla behavior
20. A juvenile walking directly towards a female while looking directly at her
   - Socially appropriate gorilla behavior
   - Socially inappropriate gorilla behavior
21. A juvenile staring at a silverback from across the habitat
   - Socially appropriate gorilla behavior
   - Socially inappropriate gorilla behavior
22. A juvenile walking past a female while looking at her and then looking away
   - Socially appropriate gorilla behavior
   - Socially inappropriate gorilla behavior
23. A silverback exploring the rear area of the habitat out of view of the other gorillas
   - Socially appropriate gorilla behavior
   - Socially inappropriate gorilla behavior
24. A juvenile staring at a female while standing 10 feet behind her
   - Socially appropriate gorilla behavior
   - Socially inappropriate gorilla behavior
25. In a gorilla group consisting of an old silverback, an old female, a young adult female, and a juvenile male, which is the least dominant gorilla?
   - The silverback
   - The older female
   - The younger female
   - The juvenile
1. The gorilla pictured standing above is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant

2. The gorilla pictured sitting above is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant
3. The gorillas pictured walking above are:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant

4. The gorilla pictured above is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant
5. The gorilla pictured above is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant

6. The gorilla pictured walking above is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant
7. The gorilla pictured walking above is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant

8. The gorilla pictured being attacked above is a:
   - Silverback
   - Blackback
   - Female
   - Juvenile
   - Infant
9. The gorilla pictured climbing above is a:
   - Silverback
   - Backtrack
   - Female
   - Juvenile
   - Infant

10. The gorilla pictured above on the left is a:
    - Silverback
    - Blackback
    - Female
    - Juvenile
    - Infant
For the next series of questions, click on the speaker icon to hear the sound for the question below the icon.

11. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear

12. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear

13. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear

14. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear
15. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear

16. The mood represented by this vocalization is:
   - Contentment
   - Annoyance
   - Anger
   - Submission
   - Fear

17. Rocks are placed in the habitat:
   - To give the gorillas something to lean against
   - To provide a source of warmth in the winter
   - To give gorillas a place to "display" to gorillas in other habitats
   - A barrier that doesn't obstruct one's view
   - To give gorillas an easily remembered place to bury food for later

   For each of the following activities or actions, is the action generally appropriate or inappropriate gorilla behavior?

18. A juvenile gazing fixedly at a female while standing 15 feet in front of her
   - Socially appropriate gorilla behavior
   - Socially inappropriate gorilla behavior

19. A female walking within one foot of a silverback
   - Socially appropriate gorilla behavior
   - Socially inappropriate gorilla behavior
20. A silverback walking directly towards a juvenile while looking directly at him
   a. Socially appropriate gorilla behavior
   b. Socially inappropriate gorilla behavior

21. A female staring at a silverback while standing 10 feet behind him:
   a. Socially appropriate gorilla behavior
   b. Socially inappropriate gorilla behavior

22. A juvenile climbing on a big rock in the habitat
   a. Socially appropriate gorilla behavior
   b. Socially inappropriate gorilla behavior

23. A juvenile walking past a silverback and looking at him and then looking away
   a. Socially appropriate gorilla behavior
   b. Socially inappropriate gorilla behavior

24. A silverback staring at a female from across the habitat
   a. Socially appropriate gorilla behavior
   b. Socially inappropriate gorilla behavior

25. In a gorilla group consisting of an old silverback, an old female, a young adult female, and a juvenile male, which is the most dominant gorilla?
   a. The silverback
   b. The older female
   c. The younger female
   d. The juvenile

108
Questionnaire

Again, thank you for participating in this experiment. To aid in interpreting the results and to get your feedback, we would ask that you fill out the following questionnaire.

Debriefing Survey:
Please circle the answer that most accurately represents your situation.

1. I have visited the gorilla exhibit at Zoo Atlanta before.
   - Never
   - Once
   - Twice
   - Three times
   - More than 3 times

2. I have interacted with other virtual environments. (If so, please list them)
   - Never
   - Once
   - Twice
   - Three times
   - More than 3 times

3. I have played first person point-of-view games (such as Quake, Unreal, or Half-Life) before.
   - Never
   - Tried them once
   - A few times
   - Several times
   - They are my favorites

4. I prefer games where I play against other people to those where I play against the computer.
   - Definitely not
   - Usually not
   - Depends on game
   - Usually so
   - Definitely so

If the experiment involved using a head mounted display and interacting with a virtual world, please answer the following additional questions.

1. The virtual world felt very real to me.
   - Not at all
   - Once or twice
   - Occasionally
   - Most of the time
   - I felt like I was there!

2. I think I learned from my interactions with the virtual environment.
   - Not at all
   - Very little
   - Some

3. I have had previous experience using a head-mounted display.
   - Never
   - Once
   - Twice
   - Three times
   - Four or more times

109
1. You are now standing in the Interpretive Center. To walk in the direction you are looking, use the button under your finger. To back up, use the button under your thumb. After getting used to the HMD and how to move around, try walking through the glass and out into the gorilla habitat.

2. You are now a juvenile gorilla, and are expected to behave appropriately.

3. You have been too disruptive and have been removed from your gorilla group. After a suitable isolation period, you will be given another chance with a new gorilla group.

4. Moats are used to separate gorilla groups from each other and from visitors without blocking lines of sight.

5. Dead trees are provided for gorillas to play with and climb on.

6. Rocks are provided for gorillas to climb on, and to display to other gorillas from.

7. Contented male.


9. You have annoyed the male gorilla by either getting too close to him or staring at him for too long.

10. You have annoyed the female gorilla by either getting too close to her or staring at her for too long.

11. The male is annoyed at another male for being too close to him.

12. The male is annoyed at a female for being too close to him.

13. The female is annoyed at another female for being too close to her.

14. Angry male gorilla! Look away and run away quickly!

15. Angry female gorilla! Look away and run away quickly!

16. The male gorilla is angry at another male gorilla.

17. The male gorilla is angry at a female gorilla.

18. The female gorilla is angry at another female gorilla.

19. The male gorilla is showing his annoyance at you by using coughing and gaze aversion.

20. The female gorilla is showing her annoyance at you by coughing and gaze aversion.

21. The male gorilla is showing his anger at you by bluff charging and beating his chest.

22. The female gorilla is showing her anger at you by bluff charging and beating her chest.

23. Gorillas relate to each other using a dominance hierarchy. At the top are the silverbacks, then the males, then the females, and finally the juveniles are at the bottom.

24. Male silverback gorilla.

25. Male blackback gorilla.

26. Female gorilla.

27. Juvenile gorilla.
# Appendix C

## Personal Space Settings

<table>
<thead>
<tr>
<th></th>
<th>Silverback</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front personal space radius</td>
<td>3.5 meters</td>
<td>2.6 meters</td>
</tr>
<tr>
<td>±(0°-45°)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Side personal space radius</td>
<td>2.6 meters</td>
<td>2.0 meters</td>
</tr>
<tr>
<td>±(45°-135°)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rear personal space radius</td>
<td>2.0 meters</td>
<td>1.4 meters</td>
</tr>
<tr>
<td>±(135°-180°)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staring personal space radius</td>
<td>9.0 meters</td>
<td>6.5 meters</td>
</tr>
<tr>
<td>Length of time to be stared at before becoming annoyed</td>
<td>5 seconds</td>
<td>10 seconds</td>
</tr>
<tr>
<td>Length of time spent annoyed before coughing &amp; gaze aversion</td>
<td>2.5 seconds</td>
<td>5 seconds</td>
</tr>
<tr>
<td>Length of time spent coughing before becoming angry</td>
<td>5 seconds</td>
<td>5 seconds</td>
</tr>
<tr>
<td>Length of time spent annoyed after annoyance cause disappears</td>
<td>15 seconds</td>
<td>10 seconds</td>
</tr>
<tr>
<td>Field of view in which staring gorillas are noticed</td>
<td>±(0°-90°)</td>
<td>±(0°-90°)</td>
</tr>
</tbody>
</table>


Jampel, Barbara (1981). "National Geographic Video: Gorilla", A National Geographic Society Special produced by the National Geographic Society and WQED/Pittsburgh.


VITA
DONALD LEE ALLISON JR.

Donald Lee Allison Junior, known to his friends as Don, was born on the ninth of August in the year 1953 in Burlington, Vermont. Son of a preacher turned teacher, Don spent his childhood living in Vermont, North Carolina, Kentucky, and Alabama. Graduating from Tuscaloosa High School in 1971, he began studying math, physics, and computer science at the University of Alabama, attending concurrently with his father, who was finishing a Ph.D. in physics at the time. His college education was interrupted by his country, which felt that it needed him for the undeclared war in Vietnam.

Don spent four years in the U.S. Air Force as a ground radio equipment repairman, and was honorably discharged as a staff sergeant in 1976. Returning to school, Don completed a B.S. degree at Bethany Nazarene College in central Oklahoma, with a double major in mathematics and physics. Matriculating to the University of Illinois, Don entered the Ph.D. program in mathematics there in 1979. While there, he also served as a teaching assistant, teaching classes in college algebra and business calculus. By 1981, Don had decided that his interests lay more in the field of computer science rather than abstract mathematics. He petitioned for and received an M.S. in mathematics, and he applied for and was accepted in the Ph.D. program in computer science at the University of Illinois. However, finances were becoming a concern, so he also tested the job market by applying at AT&T and HP. Both places offered him permanent employment.

Accepting the position at Hewlett-Packard, Don moves to Colorado Springs where he spent the next ten years working on firmware and software for HP’s line of digitizing oscilloscopes. While there, Don took video courses through National Technological University’s satellite-based distance learning program. These courses
were offered by institutions such as Northeastern University, University of Minnesota, University of Massachusetts at Amherst, and others, under the auspices of NTU, which handled the paperwork. In 1989 he met the requirements and was awarded an M.S. degree in computer engineering through NTU.

The teaching experience at the University of Illinois continued to linger in the back of Don's mind, though. Finally, taking advantage of one of HP's downsizing programs, Don enrolled in Georgia Tech to pursue a Ph.D. degree in computer science so that he could teach computer science at the college level. At Georgia Tech, Don's research interests were in computer graphics and artificial intelligence, two interests that converged in his work in the field of virtual reality. While at Georgia Tech, Don implemented the virtual gorilla environment, a virtual environment that teaches middle school children about gorilla behaviors and social interactions. This project has been the subject of extensive coverage in the national and international press and has led to the publication of several refereed papers. Currently a version of this system is installed at Zoo Atlanta where it is used to augment their educational programs.

Graduating from Georgia Tech in 2003, Don is currently employed at SUNY Oneonta College, where he is an assistant professor of computer science in the mathematical sciences department. There, he offers courses in computer graphics, virtual reality, and artificial intelligence, as well as teaching the more traditional computer science courses, and pursues research projects in virtual reality with his students and other faculty members.