AN INVESTIGATION OF THE USE OF INSTRUCTIONAL SIMULATIONS IN THE CLASSROOM AS A METHODOLOGY FOR PROMOTING TRANSFER, ENGAGEMENT AND MOTIVATION

Leslie Matthew Lunce, B.A., M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2007

APPROVED:

Jon J. Young, Major Professor
Mark H. Mortensen, Co-chair
Scott W. Warren, Committee Member
Bambi L. Bailey, Committee Member
Robin K. Henson, Chair of the Department of Technology and Cognition
M. Jean Keller, Dean of the College of Education
Sandra L. Terrell, Dean of the Robert B. Toulouse School of Graduate Studies
Lunce, Leslie Matthew, *An investigation of the use of instructional simulations in the classroom as a methodology for promoting transfer, engagement and motivation.*

Doctor of Philosophy (Educational Computing), August 2007, 205 pp., 13 tables, 3 figures, references, 179 titles.

Innovative educators seek technologies to facilitate or enhance the learning experience while taking nothing away from the message of instruction. Simulations have been shown to meet this requirement. While simulations cannot replace the teacher or the message of instruction, they can provide a deeper and more cognitively engaging learning experience.

Classroom use of simulations has been ongoing since the 1960’s. However, substantive research on their efficacy remains limited. What research has been conducted indicates that simulations possess great potential as aids to instruction. The author of this dissertation pursued this question focusing on whether simulations contribute to instruction by facilitating *transfer*, improved motivation and increased engagement.

This dissertation documents a study in which instructional simulations were used in undergraduate science courses to promote engagement, transfer and knowledge-seeking behavior. The study took place at Midwestern State University (MSU), a public university located in north-central Texas with a student population of approximately 5,500. The study ran during the fall 2006 and spring 2007 terms. Samples consisted of students enrolled in GNSC 1104 Life / Earth Science during the fall term and GNSC 1204 Physical Science during the spring term. Both courses were offered through the Department of Science and Mathematics at MSU. Both courses were taught by the
same professor and are part of the core curriculum for undergraduates in the West College of Education at MSU. GNSC 1104 and GNSC 1204 yielded samples of \( n = 68 \) and \( n = 78 \) respectively. A simulation focusing on earthquakes was incorporated into the curriculum in GNSC 1104 while a simulation which presented concepts from wave propagation was included in GNSC 1204.

Statistical results from this study were mixed. Nevertheless, studies of this type are warranted to gain a more complete understanding of how students are impacted by their interactions with simulations as well as the role simulations can play in the curriculum.
ACKNOWLEDGEMENTS

First and foremost I would like to thank Dr. Mark H. Mortensen, Ph.D. for his guidance, unwavering encouragement and patience throughout this project. Through his commitment to my success Dr. Mortensen has demonstrated what it means to be a master teacher.

I would like to think Dr. Bambi Bailey, Ph.D., Associate Professor of Education and Reading, West College of Education, Midwestern State University (MSU), without whose sponsorship, patience, participation and assistance this project would have been impossible.

I would like to think my wife, Carol S. Lunce, Associate Library Director, MSU, for her patience and support. In additional, I would like to thank her for her invaluable reference and proof-reading skills.

I would like to thank Dr. Clara Latham, Ph.D., Library Director, MSU, for generously extending library services and facilities to me throughout this project.

Others deserving thanks include, Dr. Steve Lunce, Ph.D. for much sound advice, Ms. Mitzi Lewis, ABD/Ph.D., Assistant Director for Institutional Research, MSU, for proof-reading and assistance with statistical analysis, Mr. David Wright for ever-patient and professional assistance with the Polycom videoconferencing equipment, Ms. Sherry Taylor, Disability and Rehabilitation Services Counselor, for helping me obtain recorded books and assistive technologies and Ms. Cindy Trussell, Administrative Assistant, Department of Computer Education and Cognitive Systems, College of Education, University of North Texas, for much professional and timely assistance. For their
continued support and advice I also wish to thank Ms. Mary Dzorny, ABD/Ph.D., Ms. Mary Jo Dondlinger, ABD/Ph.D. and Dr. Theresa Overall, Ph.D.

DEDICATION

This dissertation is dedicated to the memory of my elder brother, Dr. Stephen Edward Lunce, Ph.D. (1947-2007). Steve was a brother, a colleague, a mentor and a friend. This dissertation is also dedicated to my father, Carroll Edward Lunce (1912-1997). My father instilled in me at an early age a love of books, a fascination with learning and an appreciation of the value and importance of reading.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ............................................................... iii

LIST OF TABLES ................................................................. viii

LIST OF FIGURES ............................................................... ix

CHAPTER 1. INTRODUCTION ................................................... 1
  Research Problem Statement
  Purpose of the Study
  Significance of the Study
  Theoretical Foundations
  Organization of the Study
  Research Questions
  Hypothesis Tested
  Anticipated Findings
  Limitations of the Study
  Concept Definitions
    Deep Structure
    Engagement
    Fidelity
    Instructional Simulations
    Surface Structure

CHAPTER 2. LITERATURE REVIEW ........................................... 16
  Simulation-games or Gaming-simulations
  Why Simulations in the Classroom?
  Research from the Literature
    Language Instruction
    Transfer of Mathematics Skills
    Teaching Probability with Simulation
    Astronomy and Geology Simulations for
      Middle School Earth Science
    River Ecosystem Simulation for
      Middle School Earth Science
    Medical Training Simulations
    Simulations for Physics Instruction
    Software Engineering Simulation
    VR-enhanced Simulation for
      Computer Graphics Instruction
      Java Applet-based Microworlds
    Simulations for Distance Learning
    Microelectronics Instruction Simulation
    Environmental Science Instructional Simulation
    Integrating Simulations Into Homework Assignments
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SAT-ACT Score Comparisons</td>
<td>152</td>
</tr>
<tr>
<td>2</td>
<td>MSU Enrollment Data, Fall 2006</td>
<td>155</td>
</tr>
<tr>
<td>3</td>
<td>MSU Enrollment Data, Spring 2007</td>
<td>156</td>
</tr>
<tr>
<td>4</td>
<td>Student Population Data, Texas</td>
<td>157</td>
</tr>
<tr>
<td>5</td>
<td>Student Population Data, Outside Texas</td>
<td>158</td>
</tr>
<tr>
<td>6</td>
<td>Test of Homogeneity of Variance</td>
<td>166</td>
</tr>
<tr>
<td>7</td>
<td>Covariance and Pearson $r$</td>
<td>167</td>
</tr>
<tr>
<td>8</td>
<td>ANOVA Model Summary</td>
<td>169</td>
</tr>
<tr>
<td>9</td>
<td>Correlations With Course Grade</td>
<td>170</td>
</tr>
<tr>
<td>10</td>
<td>Significant Predictor Variables</td>
<td>172</td>
</tr>
<tr>
<td>11</td>
<td>Repeated Measures ANOVA</td>
<td>173</td>
</tr>
<tr>
<td>12</td>
<td>Test of Within-Subjects Effects</td>
<td>174</td>
</tr>
<tr>
<td>13</td>
<td>Test of Between-Subjects Effects</td>
<td>174</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1 Dale’s cone of experience. ......................................................... 3
Figure 2 Confidence intervals for 6 predictors. ................................. 177
Figure 3 Dendrogram. ................................................................. 178
CHAPTER 1

INTRODUCTION

Since Socrates (ca. 469 to ca. 399 BC) introduced his students to the methodology which bears his name, enlightened educators have sought new and innovative ways to present their message and engage their students (Frost, 1942; Sagan, 1979). With every new theory, methodology and technology, those that have shown success as well as those that have not, educators have worked to improve the communication between teacher and learner. Education is, above all else, a system of communication (Simonson, Smaldino, Albright and Zvacek, 2003). The instructional content is the message carried across a communication channel from the instructor to the learner (Simonson et al., 2003). Feedback is carried back along a parallel channel from the learner to the instructor (Gagné, Briggs and Wager, 1992). The result is a continues cycle of communication between teacher and student. The crucial element in this communication exchange is the message, the educational content (Dewey, 1916). Much confirmatory research has established that it is the quality and content of the message which facilitate education, not the medium by which the message is conveyed (Simonson et al., 2003). Just as technology cannot replace the teacher, technology cannot substitute for the message. Instructional content poorly constituted or conveyed without thought to the characteristics of the learner effectively adds noise to the communication channel (Simonson et al., 2003). Indeed, technology may only exacerbate and further misrepresent an already turgid message. Bearing these caveats in mind, innovative educators continue to investigate technologies which can facilitate or enhance the learning experience while taking nothing away from the message.
Instructional simulations have been shown to address these requirements (de Jong, 1991; Kelsey, 2002; Standen and Herrington, 1996; Thomas and Milligan, 2004; Zacharia and Anderson, 2003).

The use of instructional aids and devices is by no means a new concept. From biographies and contemporary writings it is known that Pythagoras of Samos (ca. 570 to ca. 490 BC) used triangular ceramic tiles to help his students master the concepts of geometry (Bronowski, 1973; Sagan, 1980). From ceramic tiles to teaching machines endowed with artificial intelligence, educators have conducted research towards developing more effective modalities to convey their content. Instructional simulations are but one of the most recently forged links in this long chain of theory, research and innovation.

Simulations cannot replace the teacher or the message of instruction. What they can do is provide a deeper and more cognitively engaging learning experience (Duffy and Cunningham, 1996). When placed with other learning modalities in Dale’s cone of experience, instructional simulations are second only to “direct purposeful experience” (Simonson et al., 2003). See Figure 1. Whether the experiences they provide are dramatized or contrived, instructional simulations can provide a qualitative, engaging environment representative of reality (Alessi and Trollip, 2001; Simonson et al., 2003). Instructional simulations can permit the learner to address ill-defined real-world problems in a physically safe, non-threatening setting (Alessi and Trollip, 2001). They can allow the learner to explore realistic, open-ended scenarios providing situated learning experiences without leaving the classroom (Alessi and Trollip, 2001).
Figure 1. Dale’s cone of experience, adapted from Wiman and Meierhenry, (1969).

When placed with other learning modalities in Dale’s cone of experience, instructional simulations are second only to real-world experience.
Instructional simulations can provide a depth of learning experience unlike any other instructional modality and can help students learn from their own mistakes (Leemkuil, de Jong, de Hoog, and Christoph, 2003).

Research Problem Statement

The first successful classroom use of an instructional simulation took place in the Harvard Business School in 1909 (Saunders, 1997). However, classroom use of simulations remained somewhat limited until the 1960s. Research on the effectiveness of simulations in the classroom as aids to instruction remains limited and to some degree problematic (Washbush and Gosen, 2001). Much extant research suffers from poor research designs, limited sample size, inappropriate data analysis techniques or lack of informative data reporting (Akpan, 2001).

Purpose of the Study

The purpose of this study was to investigate the learning effectiveness of integrating an instructional simulation into a traditional university course curriculum. The study venue was one section of an undergraduate general-science course for non-sciences majors. The instructional simulation introduced students to concepts from Earth-science, chemistry and physics. The independent variable was the presence or absence of an instructional simulation as an integral part of one unit of course curriculum. To measure the effect of the independent variable, pretest-posttest scores were collected during the treatment and control phases of the study and compared using accepted statistical methods.
Significance of the Study

Although classroom use of instructional simulations has been ongoing since the 1960s, substantive research on the topic remains limited. This is unfortunate because what research has been conducted indicates that simulations can possess great potential as aids to instruction (Aldrich, 2005; Forinash and Wisman, September, 2001). The focus of this study was to further document the significant role simulations can play in the classroom as providers of opportunities for situated learning experiences.

Further studies of this type are warranted in order to gain a more complete understanding of how students are impacted by their interactions with simulations as well as the role simulations can play in the curriculum. The questions addressed by this study focused on whether simulations contribute to instruction by facilitating transfer, improved motivation and increased engagement.

Theoretical Foundations

The fundamental conceptual elements of experience-based learning can be found in the writings of Aristotle and Socrates, subsequently refined and systematized by Dewey (Ruben, 1999). The theoretical foundation of this study was built upon the concepts of experiential learning, constructivism, behaviorism, situated learning and Vygotsky’s zone of proximal development (ZPD) (Litowitz, 1993). The importance and effectiveness of experiential learning was first documented by Piaget (1929), Dewey (1938) and Ausubel (1968). Constructivism asserts that reality is subjective, that all knowledge is derived from an individual’s actions upon the world around him/her (Gardner, 1985; von Glasersfeld, 1996). Simulations can be excellent constructivist learning tools because they present the learner with an open-ended, problem-based,
discovery-learning experience in which the learner is free to construct new knowledge.

In contrast to constructivism, behaviorism posits an objective environment in which the learner reacts to external stimulus (Sternberg, 1999). Simulations essentially present the student with a microworld containing a finite number of variables. By observing the student’s response to these variables and adjusting their values without the student’s knowledge, the instructor or tutor can modify or direct the student’s behaviors.

An outgrowth of constructivist theory, situated learning is based on the concept that learning takes places in a specific context and that context impacts the learning experience (Land and Hannafin, 2000; Rystedt and Lindwall, 2004). In the classroom students learn to deal with knowledge presented outside or divorced from the context in which that knowledge was originally encountered (Duffy and Cunningham, 1996; Gardner, 1985). Knowledge conveyed in the classroom tends to be situated in the context of the classroom and the school rather than the context in which the knowledge was created (Henning, 1998). This dichotomy has been shown to negatively impact the learning experience, adversely effecting learner motivation in particular (Henning, 1998). At the same time, real-world learning situated in real-world contexts has been shown to have positive impacts on learning and learner motivation (Duffy and Cunningham, 1996). Instructional simulations have been shown to provide a situated learning experience in the classroom by presenting knowledge in a context approaching that of a real-world setting (Engeström, 1999; Kirshner and Whitson, 1997; Lave and Wenger, 1991; Lemke, 1997).

Vygotsky’s ZPD represents the difference between a student’s demonstrated
achievement and what the student can potentially achieve given appropriate scaffolding and coaching (Litowitz, 1993). Instructional simulations are excellent examples of the ZPD concept because they require the student to stretch beyond what he/she knows.

Organization of the Study

During the early summer of 2006 I met with an Associate Professor of Education and Reading from the West College of Education at Midwestern State University (MSU) to select an appropriate venue for the study. A pilot study was conducted during the fall 2006 semester to test data collection methods and to correct procedural flaws (Gall, Gall and Borg, 2003; Isaac and Michael, 1997). The pilot study focused on one section of GNSC 1104 Life/Earth Science, an undergraduate course offered by the Department of Mathematics and Science at MSU. The course was taught by an Associate Professor of Education and Reading and is a required core course for all undergraduates in the West College of Education at MSU. Course enrollment provided a convenience sample of \( n = 68 \).

Scores were collected from content-specific pretests and posttests administered by the Associate Professor. A pretest was administered during the first half of the semester to measure student’s pre-existing knowledge of Earth-science. Students then completed a unit on geology in which the Virtual Earthquake® simulation (Desharnais, Novak, Mayo, Risner and Vasconcelos, Virtual courseware: Earthquake, California State University, Los Angeles, CA, www.sciencecourseware.org/eec/Earthquake/) served as the dominant element in the curriculum. A posttest was administered after students completed the simulation. The geology unit incorporated three chapters of
course material. These curriculum activities which centered on the Virtual Earthquake© simulation constituted the treatment phase of the pilot study.

During the second half of the semester a second content-specific pretest was administered to gage existing student knowledge of biological sciences. Students then utilized a textbook-based curriculum to complete a unit on ecology. A content-specific posttest was administered at the end of the unit. Once again, a unit included three chapters of course content. This part of the curriculum represented the control component of the pilot study. Students were also required to complete a lab exercise and daily quizzes through WebCT® online proprietary virtual learning environment (Blackboard Inc., Washington DC, www.webct.com). These scores were collected as documentation of attendance and as indicators of academic performance throughout the semester. At the end of term students were asked to complete a brief questionnaire on knowledge-seeking behaviors. All data were collated and analyzed using the Statistical Package for the Social Sciences® (SPSS) (SPSS Inc., Chicago, IL, www.spss.com). See Appendix H for a complete discussion of the pilot study.

The principal study was conducted during the spring 2007 term at MSU. The study focused on one section of GNSC 1204 Physical Science, a course in physical sciences for non science majors. The course was taught by an Associate Professor of Education and Reading and provided a convenience sample size of \( n = 78 \). Scores were collected from four content-specific pretest-posttest scales administered by the Associate Professor.

The first pretest-posttest was administered early in the semester to measure student’s pre-existing knowledge of motion and force. Shortly thereafter, a second
pretest was administered to measure students knowledge of sound waves. Following this pretest students interacted with the Properties of Mechanical Waves® simulation (Knight, San Francisco, CA : Addison-Wesley, wps.aw.com/aw_knight_physics_1/0,8722,1123696-,00.html), a physics simulation focusing on sound wave propagation. A posttest was administered after students completed the simulation. Curriculum activities centered on the Properties of Mechanical Waves simulation constituted the treatment phase of the study. Interaction with the simulation was followed by a posttest.

A third content-specific pretest-posttest focusing on electricity was administered as part of the control. Late in the semester a fourth pretest-posttest was administered focusing on the periodic table. The first, third and fourth pretest-posttest cycles represented the control component of the study.

Students were also be required to complete a lab exercise and daily quizzes through WebCT. These scores were collected as documentation of attendance and as indicators of academic performance throughout the semester.

Research Question

Can knowledge and skills acquired in an instructional simulation be transferred to a similar but non-identical setting outside the simulation?

Hypothesis Tested

\[ H_0 : \text{Transfer of knowledge and skills from an instructional setting in which simulations are used to a similar instructional setting where simulations are not used cannot be demonstrated to have occurred.} \]
$H_1$: Transfer of knowledge and skills from an instructional setting in which simulations are used to a similar instructional setting where simulations are not used can be demonstrated to have occurred.

Anticipated Findings

It was anticipated that this study would demonstrate that simulations provide improvements to the classroom learning experience. Students would experience knowledge growth as demonstrated by transfer of skills and concepts mastered through interaction with a simulation to a similar but non-identical context outside the simulation. Skills transfer and concept mastery would be demonstrated by improved test scores and knowledge-seeking behavior.

Limitations of the Study

Validity and reliability of the Virtual Earthquake simulation software used in the pilot study have been documented in the literature by Fisk, Boucher and Hanley (1998) and Novak (1999) as well as Deshamais, Novak and Mayo (2002). These criteria are based on data collected from 38 domain experts in 14 countries who used the simulation as part of their regular course curriculum (Fisk et al., 1998). Validity and reliability of the Properties of Mechanical Waves simulation used in the full study have been documented in the literature by Adams, Finkelstein, Gratny, Perkins and Wieman (2006), Sands (2004), Schroeder (2006), and by MacIsaac (WebSights, 2005). Domain experts from over 120 universities and colleges across the United States have reviewed and contributed to the text as well as used and evaluated the simulation in the classroom (Knight, 2004). Further assessment of validity and reliability of simulation
software are beyond the scope of this study.

This study did not compare instructional simulations to traditional classroom instruction as this would have constituted yet another media-comparison study (Clark, 1983). Such studies have been shown to produce little in the way of significant statistical results (Nathan and Robinson, 2001).

The pilot study focused on one section of GNSC 1104 Life/Earth Science, an undergraduate course offered by the Department of Mathematics and Science at MSU, during the 2006 fall term. This is a core course required for all undergraduates in the Wes College of Education at MSU. Students enrolled in the courses were considered a convenience sample drawn from the overall undergraduate population of the university. The course was taught by an Associate Professor of Education and Reading in the West College of Education, MSU.

The full study focused on one section of GNSC 1204 Physical Science an undergraduate course offered by the Department of Mathematics and Science at MSU during 2007 spring term. This is a core course required for all undergraduates in the West College of Education at MSU and was taught by the same Associate Professor of Education and Reading who taught GNSC 1104 the previous fall term. Students enrolled in the courses were considered a convenience sample drawn from the overall undergraduate population of the university.

The venue for the pilot study and the full study did not permit splitting the course sections into separate treatment and control groups. As a result, a pre-experimental one-group, pretest-posttest design with control was used for the full study. This limitation also precluded random assignment necessitating selection of a convenience
sample (Gall, et al., 2003). In addition, student participants were not to undergo testing or be required to complete assignments other than what they would encounter as part of the normal course curriculum. These limitations made controlling for the *Hawthorne Effect* somewhat problematic (Cook and Campbell, 1979; Gall et al., 2003; Isaac and Michael, 1997). Venue limitations also precluded administration of a personality survey instrument or conduct of debriefing activities post-simulation. Finally, participants were not offered extra credit, gifts or other incentives to participate in the study.

**Concept Definitions**

*Deep structure*

Deep structure refers to the psychological mechanisms of the simulation and the nature of interactions which take place between the simulation participants and the simulation model or scenario (Gredler, 1992). Deep structure defines the kinds of participant behaviors essential for addressing the problem or exercise. In addition, deep structure determines which participant behaviors will be reinforced as a result of interacting with the simulation scenario.

*Engagement*

Also referred to as *student engagement*, this concept is a measure of student’s motivation to fully focus, to participate, to achieve, to affectively and cognitively connect with the learning process (Quinn, 2005). Students who exhibit engagement behaviors tend to self-select tasks which challenge their competencies. These students will also initiate learning activities and expend extraordinary efforts towards the completion of learning exercises. Such students view learning as a positive experience and
demonstrate enthusiasm, optimism and curiosity throughout the learning process (Brewster and Fager, 2000). Confirmatory studies have established a significant causal correlation between high levels of student motivation and engagement and low drop-out rates (Kushman, Sieber and Heariold-Kinney, 2000).

**Fidelity**

For all categories of simulation, *fidelity, validity or reality of function* refers to the accuracy with which the simulation models a real-world system or phenomena (Alessi and Trollip, 2001; Gredler, 1992). Fidelity also refers to the realism of learner interaction facilitated by the simulation, as well as the type and frequency of feedback provided. A well designed simulation can maintain a high degree of fidelity while abstracting or omitting distracting elements that would otherwise be present in a real-world situation (Alessi and Trollip, 2001; Moore, Burtan, and Myers, 1996). Simulations created with the goal of supporting fidelity or reality of function require attention to three specific design aspects (Gredler, 1992). First, all rules for operating within the simulation environment must be authentic and as realistic in nature as what the participant would confront in a real-world setting. Second, sufficient documentation must be provided to the participant to facilitate gaining a grasp of the key elements of the problem to be confronted. Any such documentation should be formatted as memos, email messages, faxes, newspaper clippings, news reports or other real-world information sources. Third, the design of the simulation should support consistent behaviors and conscientious reactions on the parts of the participants. Rieber (1994) has presented research in support of a nonlinear relationship between learning and a simulation's level of fidelity. This relationship is to a large degree depends on the instructional level of the student.
Accordingly, increasing the fidelity of a simulation will not in and of itself improve the learning experience. The type of documentation and feedback provided to the simulation participant are also dependent on whether the simulation is defined as instructor supported or stand-alone, and whether it supports dynamic learning or linear learning (Aldrich, 2005).

In de Jong, and van Joolingen (1998) and Morgan Vista, (2005) fidelity has been further categorized as perceptual, manipulative or functional. Perceptual fidelity refers to how closely the simulation models the content being taught. Manipulative fidelity refers to the degree to which learner's actions in the simulation correspond to similar actions to be taken in a real-world scenario. Functional fidelity represents a measure of how well the simulation model and the simulation “deep structure” correspond with the structure of the material being taught (Gredler, 1992). Regardless of the type of fidelity, all simulations are simplifications of reality with the result that fidelity is reduced from what the learner would experience in real-life. The degree of fidelity in the simulation should be determined by instructional effectiveness rather than a designer’s desire to precisely simulate a real-world situation (Alessi and Trollip, 2001).

*Instructional Simulation*

Lorenz, Horstmann, Oesker, Gorczytza, Dieckmann and Egelhaaf (2004) presented a definition of instructional or educational simulations which differentiates them from other types of simulations. This definition is based on three factors. First, instructional simulations incorporate an instructional design, present the learner with concrete tasks, include a glossary, references and other similar resources, as well as a note taking facility for working out solutions. Second, the overall design of the simulation
permits both synchronous and asynchronous manipulation of variables and inputs. Third, animation and interactivity are used to render the simulation interface spatially and temporally dynamic.

**Surface Structure**

Surface structure as defined in Gredler (1992, 2001) constitutes the paraphernalia, user interface and observable mechanics of a simulation or game. Surface structure also refers to the dynamics of feedback, problem presentation and other interactions between the simulation and the participant. The scenario and associated data to be addressed by the participant represent the essential surface structure components in a simulation.
Alessi and Trollip, (2001) have defined an instructional simulation as “a model of some phenomena or activity that users learn about through interaction with the simulation”. These authors exclude movies, animation and some types of games from their definition as these formats are not based on an internal model. Instructional simulations have been broadly categorized by Gredler (1998) as either experiential or symbolic. The goal of experiential simulations is to situate the learner in a clearly defined role as part of a psychological reality established within the simulation, and to engage the learner in guided learning-by-doing activities (Debe, 1996). Symbolic simulations dynamically represent the behavior of a real-world system, processes or phenomena. The behavior is simulated through the interaction of two or more variables. By manipulating these variables the learner can engage in theory building, hypothesis testing and discovery-learning. Aldrich (2005) has defined four broad genres or instructional content models of simulations. These genres are branching stories, interactive spreadsheets, game-based models, and virtual labs or virtual projects. All simulations embody five specific characteristics.

1. Simulations are problem-based learning environments in which the learning experience is initiated by some perceptive task or event.
2. Participants are confronted with resolving a problem in a real-world setting with limited information, limited time and other constraints.
3. Simulation problem scenarios tend to be open-ended or poorly defined with outcomes not determined by luck or chance. Participants are expected to
work through the problem scenario making the best decisions they can while learning from their mistakes.

4. Most participants in simulations carry out functions associated with roles assigned to each participant before entering the simulation.

5. Participants experience reality of function provided they carry out their assigned roles effectively and conscientiously (Gredler, 1992).

As instructional exercises, simulations can motivate the learner to engage in problem solving, hypothesis testing, experiential learning, schema construction, and development of mental models (Chong, 1999; Garris, Ahlers and Driskell, 2002). To achieve these goals, three components of the simulation are essential (Quinn and Conner, 2005).

1. The simulation must be relevant and authentic, and should begin with some precipitating action or event which motivates the participants to take action (Aldrich, 2004). For example, the student may be confronted with a regional disaster in which flooding is occurring in a densely populated coastal area.

2. Roles for simulation participants are defined and structured to interact with various key aspects of the problem scenario. Participants must take an active part in the scenario. Their roles will not permit them to stand back and observe while others take action. The scenario should apply emotional pressure to the participants forcing them to act (Aldrich, 2004). For example, medical personnel engaged in a combat field simulation cannot hold back to observe how the scenario will unfold. They must take up their roles just as they would in a real-world combat situation.
3. Participants operate in a realistic situation in which the consequences of their actions become quickly apparent. As a result, participants address problems in the simulation scenario seriously and conscientiously. At the same time, the scenario should provide the participants with a sense of unrestricted options (Aldrich, 2004). An example could be medical students questioning a potential patient. In order to acquire usable experience from the simulation, the students must embed the simulation with as much realism as possible which means thinking of the situation as real.

To facilitate learning, educational simulations rely heavily on scaffolded practice, coaching, feedback and reflection (Duffy and Cunningham, 1996: Quinn and Conner, 2005; Standen and Herrington, 1996).

A simulation is based on an internal model of a real-world system or phenomena in which some elements have been simplified or omitted in order to facilitate learning (Aldrich, 2005). These models may be generally characterized as glass box or black box (Thomas and Milligan, 2004). The glass box designation indicates a situation in which the student has access to and can modify the simulation model whereas the model is hidden from the student in a black box simulation (Alessi, 2000). Simulations can be further characterized as qualitative or quantitative as in de Jong and van Joolingen (1998). Quantitative simulation models are usually derived from a set of mathematical functions. Qualitative simulation models are usually derived from a set of heuristics or rules. Qualitative simulations can therefore be more flexible and dynamic than qualitative simulations. As presented in Gredler (1992), theoretical models upon
which simulations are built tend to be of three general types: continuous, discrete or logical.

1. *Continuous models* are constructed using calculus in order to represent a system with an infinite number of states.

2. *Discrete models* employ statistics and queuing theory to represent systems with quantitatively discrete states.

3. *Logical models* are most often represented using a set of heuristics implemented through a high-level computer programming language.

Logical models are utilized most often in educational simulations. Continuous and discrete models are most often found in scientific and engineering simulations. An educational game should not be considered a simulation unless the game is built upon a model of a real-world system or phenomena (Alessi and Trollip, 2001). Modeling an event, procedure or situation does not in itself render the model a simulation.

Simulations are generally grouped into four categories: physical, iterative, procedural or situational as documented by Gredler (1992).

1. Physical simulations allow the learner to manipulate variables in a open-ended scenario and observe the results. The learner may also manipulate physical objects which are designed to closely replicate real-world objects in form and function. An example of a physical simulation would be a model of global weather patterns in which the student can manipulate certain parameters and observe the outcome.

2. Iterative simulations tend to focus on discovery learning by providing the student with opportunities to conduct scientific research, build and test
hypothesis and observe the results. This type of simulation typically focuses on teaching phenomena which are not readily observable in real-time, for example, phenomena from biology, geology or economics. In this case, the student would repeatedly run the simulation, altering variables with each iteration to test a hypothesis.

3. Procedural simulations require the student to manipulate simulated objects with the goal of mastering the skills required to correctly and accurately manipulate physical objects in a real-world setting. A typical example of a procedural simulation is a chemistry lab experiment in which the student manipulates simulated laboratory equipment with the goal of preparing the student for working in a real-world laboratory setting (Aldrich, 2005).

4. Situational simulations generally model human behavior focusing on attitudes of individuals or groups in specific settings. These types of simulations, also referred to as digital-decision role-play simulations, often employ role playing as a vehicle to allow students to explore different options and decision scenarios (Cohen et al., 2006). Situational simulations are usually designed to be run several times in a specific context or situation with each participant in the simulation scenario playing a different role in each iteration (Cohen et al., 2006; Wilson and Cole, 1996). It should be noted that, because of their open-ended design, and due to the complexity of modeling human behavior, situational simulations tend to be the most difficult type of simulation to design and utilize effectively.

Physical and iterative simulations are generally defined as conceptual or about
something while procedural and situational simulations are defined as operational because they focus on how to do something (Alessi, June 2000; de Jong and van Joolingen, 1998). For all categories of simulation, fidelity or reality of function refers to the accuracy with which the simulation models a real-world system or phenomena (Alessi and Trollip, 2001; Gredler, 1992). Fidelity also refers to the realism of learner interaction facilitated by the simulation, as well as the type and frequency of feedback provided. Indeed, Rosenfield et al., (2002) have documented 4 essential characteristics of effective feedback.

1. Feedback must provide the participant with a clear understanding of the gap between the current state of the simulation and the simulation goal state.
2. The simulation must maintain a record of corrective actions taken by the participant and provide confirmation of the participant’s actions.
3. Feedback must be concise and easy for the participant to understand.
4. The participant must have demonstrated mastery of the prerequisite skills, concepts and knowledge to respond correctly to feedback before interacting with the simulation.

A well designed simulation can maintain a high degree of fidelity while abstracting or omitting distracting elements that would otherwise be present in a real-world situation (Moore et al., 1996; Rieber, 2000). Simulations created with the goal of supporting fidelity or reality of function require attention to three specific design aspects (Gredler, 1992). First, all rules for operating within the simulation environment must be authentic and as realistic in nature as what the participant would confront in a real-world setting. Second, sufficient documentation must be provided to the participant to facilitate
gaining a grasp of the key elements of the problem to be confronted. Any such
documentation should be formatted as memos, emails, newspaper clippings, faxes,
news reports or other real-world information sources. Third, the design of the simulation
should support consistent behaviors and conscientious reactions on the parts of the
participants. The type of documentation and feedback provided to the simulation
participant is also dependent on whether the simulation is defined as instructor
supported or stand-alone, and whether it supports dynamic learning or linear learning
further categorized fidelity as perceptual, manipulative or functional. Perceptual fidelity
refers to how closely the simulation models the content being taught. Manipulative
fidelity refers to the degree to which learner's actions in the simulation correspond to
similar actions to be taken in a real-world scenario. Functional fidelity represents a
measure of how well the simulation model and the simulation deep structure correspond
with the structure of the material being taught (Gredler, 1992).

Simulations may be further characterized by their surface structure and deep
structure. Surface structure as defined in Gredler (1992, 2001) constitutes the
paraphernalia, user interface and observable mechanics of a simulation or game.
Surface structure also refers to the dynamics of feedback, problem presentation and
other interactions between the simulation and the participant. The scenario and
associated data to be addressed by the participant represent the essential surface
structure components in a simulation. Deep structure refers to the psychological
mechanisms of the simulation and the nature of interactions which take place between
the simulation participants and the simulation model or scenario. Deep structure defines
the kinds of participant behaviors essential for addressing the problem or exercise. In addition, deep structure determines which participant behaviors will be reinforced as a result of interacting with the simulation scenario. The goal of the underlying simulation model is to reinforce correct behaviors. Careful observation and documentation of a simulation’s deep structure can help the designer and the educator determine whether the simulation is reinforcing appropriate behaviors.

Gredler (1992) has noted two types of prior knowledge which the student must possess prerequisite to successfully interacting with an instructional simulation. First, the student must be well grounded in domain-specific knowledge and major concepts in order to understand the content and problems presented in the simulation environment. This knowledge is essential if the student is to systematically manipulate variables and properly conduct research using the simulation. Second, the student must be possessed of a systematic strategy for addressing the type of open-ended, multi-variable scenarios encountered in the simulation. Helping the student acquire this knowledge may require informing the student about the level of task complexity to be expected in the simulation, as well as the number and purpose of variables to be manipulated.

While participant interaction with a simulation typically lasts for one class meeting or perhaps a few hours, extended simulations continue over multiple class periods or days (Hertel and Millis, 2002). Extended simulations overcome the time limitations of most instructional simulations and therefore can present dozens of learning objectives. Participants have significantly increased time to collect data, explore problem scenarios, form theories, experiment, conduct analysis and become more deeply invested in their
roles. However, extended simulations tend to be most effective only with small groups of mature participants.

Simulations have a number of advantages over other instructional methodologies and media. One need only look to the United States Military services which use simulations more than any other training modality (Aldrich, 2005; Quinn and Conner, 2005). Simulations have been shown to provide a motivating and satisfying learning experience because they incorporate fundamental motivational elements, i.e. challenge, curiosity, control and fantasy (Cohen et al., 2006). Motivation cannot be imposed, it must be engendered in the learner by providing the learner with choice and a sense of control over the learning experience (Hertel and Millis, 2002). Active learning can promote motivation resulting in deep learning. Deep learning is characterized by persistent learning or retention of knowledge and skills beyond the learning environment, better understanding of concepts, and critical thinking ((Hertel and Millis, 2002; Quinn and Conner, May 2005). Simulations constitute an active learning environment which can bring about a deep learning experience for the student.

Some educators see simulations as a promising modality because they can facilitate a constructivist, learner-driven environment which encourages discovery learning and knowledge acquisition (Horiguchi and Hirashima, 2004). Students often find active participation in simulations to be more interesting, intrinsically motivating and closer to real-world experiences than more traditional classroom-based learning activities (Rieber, 2000). Simulations have been shown to provide deep learning and skills transfer with the result that what is learned facilitates improved performance in real-world settings (Hertel and Millis, 2002; Leemkuil et al., 2003). Confirmatory studies
documented by Hertel and Millis (2002) have established the learning effectiveness of simulations with non-traditional students. Further, there is evidence to suggest that simulations may be more efficient modalities for learning in some content areas (de Jong, 1991). Simulations can be very flexible in that both student and instructor can have a high degree of control over simulation variables (Duffy and Cunningham, 1996; Hung and Chen, 2002). Simulations allow students to experience phenomena which could be dangerous, expensive or even impossible to observe in the real world (Reiber, 2000). For example, simulations permit the learner to stretch or compress time and space (Wilson and Cole, 1996). Because simulations are simplifications of real-world phenomena, they facilitate learning by omitting what would otherwise be distracting elements in a real-world situation (Alessi and Trollip, 2001). Confirmatory studies documented by de Jong and van Joolingen (1998) indicate that simulations are twice as efficient as other learning modalities when efficiency is defined as time spent learning. As documented in Quinn and Conner (2005), simulations provide retention and learning affect which extends beyond the learning situation. Finally, simulations can accommodate a wide range of instructional strategies, including microworlds, scientific discovery learning, virtual reality, laboratory simulations, role playing, case-based scenarios and simulation-gaming (Dalgarno, 1996).

Simulations do have distinct disadvantages compared with other modalities. Because simulations are often used with problem-based learning methods, they stimulate learners to immerse themselves in a problematic situation and experiment with different approaches (Heinich et al., 1999). Problem-based learning has been defined as a constructed, exploratory learning environment in which student learning is
guided with careful scaffolding to achieve optimal benefit (Goldman, Zech, Biswas & Noser, July 1999). Problem-based learning has been further characterized as a process whereby knowledge from past problem solving experiences is transferred to new scenarios which share significant aspects with previously experienced problems such that transferred knowledge may be applied to the creation of new problem solutions (Reimann and Schult, 1996). This type of learning may require significantly more time than other methods of instruction. Research has shown that without appropriate coaching and scaffolding (Duffy and Cunningham, 1996), as well as appropriate feedback and debriefing (Leemkuil et al., 2003), the learner gains little from the discovery learning simulations can facilitate (Min, 2001; Rosenfield, Lou and Dedic, 2002). In addition, research has indicated that in the absence of reflection and debriefing, students tend to interact with a simulation as merely a game (Leemkuil et al., 2003). Some constructivists argue that educational simulations oversimplify the complex and diverse nuances of real-life experience giving the learner an inaccurate understanding of a real life problem or system (Amory, 2001). This observation is worthy of note in that simulations do not usually incorporate competition. In many real-world settings competition is an important motivating factor. Debate is ongoing as to whether competition has a positive impact on the learning effectiveness of simulations. In most simulation scenarios there is no explicitly correct or explicitly incorrect answer. Some students may experience a high degree of discomfort or anxiety when confronted with this type of open-ended problem (Cohen et al., 2006). Finally, development of educational simulations may involve extensive planning and require significant investment of time, labor and financial resources.
Simulation-Games or Gaming-Simulations

Aldrich (2005) has noted that while the newest video games are increasingly abstracted from simulations, simulation interfaces are becoming increasingly game-like. A number of the latest video games were initially developed as simulations modeling a specific situation or set of circumstances. The popular Maxis® SimCity (Electronic Arts Inc., Redwood City, CA, www.maxis.com) game and its various incarnations, Sid Meier’s® Civilization IV (Take-Two Interactive Software, Inc., New York, NY, www.2kgames.com/civ4/home.htm) and Microsoft® Age of Empires III (Microsoft Corporation, www.microsoft.com) are representative examples of the continuing synthesis of games and simulations (Rehberger, 2006). This trend should come as no surprise as developers of simulations and video games continue to seek a workable synthesis of the two modalities (Gredler, 1998; Leemkuil et al., 2003). One approach to achieving this goal has been to assign simulation participants, or teams, to serious rolls which place them in directly competitive situations. Upon completion of the simulation scenario, a winner is designated based on the participants’, or teams’, performance in the simulation. As a result of merging games and simulations, the two modalities are no longer distinct and separate but rather represent two ends of a continuum.

Nevertheless, research has shown that attempts to merge simulations and games can send confusing messages to the learner, resulting in a negative learning experience (Cohen et al., 2006; Gredler, 1998). Friction may also arise between those simulation participants who take the scenario seriously and those who approach it is merely a game.

Gredler (1992, 1998) has identified 3 characteristics of games which differentiate
them from simulations as well as 4 specific problems which can arise when games and simulations are combined. A game is designed to be a world unto itself with its own internal rules which have no counterpart in the real world. Competition is usually the paramount motivating factor. Second, game paraphernalia and rules for play may be extremely complex and usually are designed to permit one player or team to defeat other players or teams. Third, games are about winning by any means not explicitly proscribed by the rules. When these 3 characteristics of games are combined with the 5 typical characteristics of simulations previously noted, the outcomes can be significantly problematic. First, a participant or team that is losing may attempt to crash the simulation scenario prior to completion. Second, winning is frequently more a factor of the abilities of the most skilled member of a team rather than a synthesis of all team member skills. Third, games tend to reinforce behaviors which culminate in winning. Such behaviors are usually counterproductive to the intended learning outcomes of simulations. Finally, games are rarely designed with the type of open-ended learning environment (OLE) that simulations can provide (Alessi and Trollip, 2001). Games can provide an effective learning experience. However, their focus on participation, specific skills mastery and exploration of relationships between known concepts is quite different from that of simulations. Simulations represent a more complex learning experience focusing on exploration, creation of new knowledge and the resolution of ill-defined problems (Cohen et al., 2006). In short, games focus on rules and competition while simulations focus on rules and fidelity. Increasingly, designers and educators agree that both simulations and games have their places in the classroom. The development of multi user virtual environments (MUVEs) combining characteristics of both games and
simulations have shown promise as viable learning resources (Clarke and Dede, April, 2005; Ketelhut, Dede, Clarke and Nelson, April, 2006). Current research indicates that careful combining of elements from both simulations and games can result in viable learning experiences.

Why Simulations in the Classroom?

Users express themselves through their input; the simulation performs calculations and then presents feedback in the form of an output. Then the cycle repeats. In simulations, the speed of the input-calculation-output workings is ten, or a hundred, or even a thousand times faster than real life or the classroom’s understanding-planning-experiencing-updating cycle. (Aldrich, 2004).

Aldrich’s comments notwithstanding, it is probably impractical and even undesirable to consider transforming every classroom setting into a situated learning environment. However, it may be possible to bring at least some of the benefits of situated learning into the conventional classroom setting. Simulations may make this possible.

Simulations can provide a method for students to check their understanding of the real-world by modeling the structure and dynamics of a conceptual system or a real environment (Heinich et al., 1999). Simulations facilitate situated learning by providing interactive practice of real-world skills, focusing on the essential elements of a real problem or system (Berge, 2002; Heinich et al., 1999). Simulations can “communicate complex and technical scientific information” (Saul, 2001) allowing students to acquire tacit knowledge in a manner similar to that experienced by domain experts (Polanyi,
Tacit knowledge has been defined by Varlander (2006) as personalized knowledge which individuals possess, but which can be problematic to pass on. Successful transfer of tacit knowledge typically requires close social interaction, observation of domain experts and learning by doing. Related to tacit knowledge, reified knowledge is acquired through participation by the learner (Varlander, 2006). This participation is characterized by interaction between the learner and an expert, or members of a community of practice, which focus on negotiation, bargaining and translation. Simulations in which a novice works under the observation of an expert are well suited to providing this type of knowledge.

A well designed simulation can engage the learner in interaction by helping the learner predict the course and results of certain actions, understand why observed events occur, explore the effects of modifying preliminary conclusions, evaluate ideas, gain insight, and stimulate critical thinking. Simulations can also provide the learner with “feedback throughout the learning process” (Granland et al., 2000). Because simulations are adaptable and dynamic, they can guide the learner in the achievement of specific learning goals (Gibbons et al., 1997). It seems apparent that many benefits of situated learning can be provided to the learner in the traditional classroom through the use of simulations.

Just as each learner has a different learning style, there are many types and degrees of situated learning. Hung and Chen (2002) present situated learning as occupying a continuum of instructional contexts ranging from *authenticity to generalizability*. The authors assert that with appropriate scaffolding, simulations can serve in the mid-range of this continuum by providing real-world, problem-based
learning in the classroom. However, this functionality can be provided only when simulations allow both student and instructor the widest possible range of variables which can be set and manipulated. In so doing, simulations can approach the realities of a situated learning context.

As the instructional use of simulations in the classroom becomes more important, so too does the necessity of research into the efficacy of instructional simulations. In the academic, public and corporate sectors efforts are ongoing towards developing and evaluating the use of simulations to facilitate situated learning. Evaluation and assessment are crucial areas in need of research as Java applets and hand-held computing facilitate the deployment of simulations over the Internet (Granland et al., 2000; Osciak and Milheim, 2001).

Research from the Literature

Language Instruction

Perhaps one of the most interesting and technically challenging instructional simulations documented in the literature was reported by Harless, Zier and Duncan (1999). These authors discussed the Virtual Conversation™ project, an ongoing collaborative development effort between Interactive Drama Incorporated (IDI) of Bethesda, Maryland (see http://www.idrama.com/idihome/) and the Defense Language Institute Foreign Language Center (DLIFLC) (see http://pom-www.army.mil/pages/dliflc.htm), a department of the U. S. Army Research Institute (see http://www-ari.army.mil/). The DFILIC is part of the Defense Advanced Research Projects Agency (DARPA) (see http://www.darpa.mil). The intended audience for the Virtual Conversation™ project are military and government personnel who at some
point in their careers have achieved a Level 2 proficiency in reading, listening and speaking Modern Standard Arabic (MSA), but who need to regain their proficiency in a very short time (see http://www.dlielc.org/testing/opi_levels.html). The designers of Virtual Conversation™ believe they have made this goal achievable by developing a simulation which allows learners to “converse independently with native speakers in a multimedia environment” (Harless et al., 1999).

Although the authors do not discuss any internal details of the simulation, its primary function is to allow language learners to interact with native language speakers through role-playing and natural language dialog. The native speakers are presented through full-motion, interactive digital video of real-life native speakers. These native speakers relate both real and fictional scenarios from their culture in order to engage the language learner in realistic conversation. As the learner participates in the conversation and presents questions to the native speaker, the direction, tone and content of the dialog provided by the native speaker are dynamically altered to further simulate the flow of natural conversation. The native speaker asks questions of the learner which the learner is expected to answer entirely in MSA. The overall goal of the designers was to give the learner the feel of interacting with a live person throughout the simulation.

Virtual Conversation™ also incorporates many of the features one would expect to find in more traditional language learning software. The learner can refer to language dictionaries as well as vocabularies of questions and phrases. These materials are presented to the learner on his/her computer screen in both English and MSA. The learner may also record and listen to his/her own pronunciations comparing these to the
pronunciations of a virtual instructor who can be contacted at any point during the simulation. As the student interacts with the virtual instructor, the simulation displays a confidence level providing immediate feedback to the learner as to their achieved level of proficiency.

Students who initially tested Virtual Conversation™ were asked to repeatedly interact with the simulation for intervals of 6 to 8 hours over a 5 day period. The objective was to simulate the type of real-life interactions the student might experience when interacting with a live native speaker during the course of one day. The authors reported that most students found the simulation so engaging that they voluntarily interacted with it for more hours than were originally requested.

The designers of Virtual Conversation™ conducted a detailed quantitative study of the simulation in order to determine its effectiveness. Participants in the study were military linguists who had previously achieved a Level 2 proficiency in speaking, listening and reading MSA, but who had been away from formal language training for more than two years. In addition, these participants had no opportunities to practice their language skills with native speakers during the intervening two-year period. All participants were selected at random from a pool of student volunteers from the Middle East School of the DLIFLC. All participants were randomly assigned to one of two teams, each team supervised by two instructors. All participants were administered a pretest which consisted of a 45 minute interview conducted entirely in MSA and the standard Defense Language Proficiency Test (DLPT). The results of these measures were used as baseline scores for each participant. After the pretest, participants were given a brief orientation to the Virtual Conversation™ simulation. Each participant was
then provided a laptop computer, headphones, microphone, a journal, a copy of the simulation software and were instructed to interact with the simulation for at least 8 hours per day for 5 days. Participants were asked to record thoughts and impressions of the simulation in their journals.

All participants were subsequently administered a post-test which consisted of a 45 minute interview conducted entirely in MSA and the DLPT. Participants were also administered a Likert-based survey and participated in open-ended interviews with faculty from the Middle East School of the DLIFLC. Test scores and survey results were analyzed using a one-sample t-test and analysis of variance (ANOVA). Significance levels were set at $\alpha = .05$.

The authors reported statistically significant increases in speaking (0.80), listening (0.50) and linguistic skills (0.23) for all participants as reflected in differences between pretest and posttest scores. Participant responses to the simulation were very positive as reflected in results obtained from the survey instrument, open-ended interviews and notations from their daily journals. The most striking aspect of the experience reported by participants was the feeling of interacting with a live native speaker. In addition, participants stated that the feel of conversing with a live native speaker significantly improved their motivation to interact with the simulation. The motivational impact of simulations is well documented in studies sited by Hertel and Millis (2002). All participants in the Virtual Conversation™ study asserted that the simulation surpassed any previous language learning experience.

These results would seem to support the hypotheses stated by Harless et al. (1999) that the Virtual Conversation™ simulation can help linguists to quickly and
accurately improve or regain language proficiency. This project also demonstrates the role interactive video can play in providing a realistic learning experience with instructional simulations. In addition, the project’s success as demonstrated by participant scores seems to indicate that a well designed simulation can serve as its own assessment; “If a learner can complete it, he or she can do the task” (Aldrich, 2001). Freiermuth (2002) and García-Carbonell et al. (2001) documented similar studies in which use of instructional simulations for language instruction resulted in a positive learning experience for students.

Criminal Justice Training

Cohen et al., (2006), reported on a project to teach the basic elements of the criminal justice system. The project centered on a CD-ROM-based simulation of typical courtroom proceedings. Participants for the study were selected from undergraduate student populations at Tufts and Northwestern Universities. The sample was divided into a control (n = 36) and two treatment groups. One treatment group (n = 51) used a paper-and-pencil version of the simulation while the other treatment group (n = 30) used the CD-ROM version of the simulation. The authors reported that treatment groups performed better on content-based examinations than the control group. In addition, both treatment groups performed at equivalent levels. This second conclusion provides confirmatory evidence that it is the simulation itself which provides the learning experience, not the technology by which the simulation is delivered. A confirmatory study was reported by Cavallier in Cohen et al. (2006) in which 180 undergraduate students at Carnegie Mellon University used the Allwyn Hall CD-ROM based simulation to master conflict resolution skills. A pencil-and-paper pretest and posttest were
administered to assess the ability of students to analyze both sides of a conflict and effect a resolution. Students who completed the simulation demonstrated marked improvement in their conflict resolution skills as well as demonstrated abilities to apply conflict management techniques learned in class.

Transfer of Mathematics Skills

Van Eck and Dempsey (2002) reported on testing a simulation designed to facilitate transfer of mathematics skills from a learning context to a real-world scenario. The simulation software was developed using Macromedia® Authorware 5.1 (Adobe Systems Incorporated, San Jose, CA, www.adobe.com) for Microsoft® Windows (Microsoft Corporation, www.microsoft.com), and incorporated interactive video as a vehicle for providing contextualized advisement to learners. The simulation metaphor was a home remodeling project which required the learners to solve problems in geometry focusing on the concepts of volume and area of simple geometric shapes. As students worked through the problems presented in the simulation, they could seek advice from a virtual aunt and uncle. The student contacted the virtual aunt and uncle with a walkie-talkie which was presented as part of the simulation interface. When contacted, the aunt and uncle appeared in the simulation as interactive digital video of real-life actors composited onto the simulation environment. A set of possible enquiry topics was presented to the student in an on-screen text box. The student could select one or more of these topics. The aunt and uncle then discussed the enquiry topic by means of digital audio recordings of real actors.

An additional aspect of the simulation was the presence of a computer-generated character working in another part of the house and confronting the same geometry
problems presented to the learners. The ability level, gender and ethnicity of this computer-generated character were selected by the student at the beginning of the simulation. The character appeared in the corner of the student’s computer screen at random intervals during the problem-solving process. The character displayed facial expressions conveying concentration or puzzlement. Occasionally the character spoke brief phrases aloud such as, “I think I’ve figured it out.” The author’s stated objective in adding the virtual character to the simulation was to provide the learner with a sense of competition.

The target population for this study was 7th and 8th grade students ranging in age from 12 to 15 years. From this population, a sample of 328 students was selected from four middle schools in a Gulf Coast city of the United States: School A (n = 50), school B (n = 75), school C (n = 123), and school D (n = 80). Students were then randomly assigned to a control group or one of four treatment groups. All students were administered a 16-item survey which collected a range of demographic data including age, gender, computer experience, mathematics experience, game playing behavior, etc. A 23-item pre-test was then administered to all students to verify that all participants possessed the basic mathematic skills necessary to address the problems presented in the simulation.

All five groups were evaluated on the same problems in geometry based on mathematics curriculum standards set out by the National Council of Teachers of Mathematics (see http://www.nctm.org/). Students in the control group were given a computer-based tutorial in which geometry concepts to be mastered were presented as word problems. This tutorial was not a simulation and did not incorporate graphics,
animation or any form of contextual advisement other than straightforward help menus. Students in the four treatment groups were presented with the home remodeling simulation. The independent variables measured over the four treatment groups were as follows.

- Treatment group 1, contextual advisement with competition.
- Treatment group 2, contextual advisement without competition.
- Treatment group 3, competition without contextual advisement.
- Treatment group 4, no contextual advisement and no competition.

Treatment groups and the control group were administered a post-test in the form of a second simulation. In this instance, the metaphor for presentation of the geometry concepts was an indoor movie theater. Students were asked to solve the same types of geometry problems they had encountered in the home remodeling simulation, but in a different, more realistic context. The post-test simulation did not include elements of competition or contextual advisement. Upon completion of the post-test, a 10-item questionnaire was administered to all students to acquire qualitative data regarding learning modality preferences and student attitudes about various aspects of the simulation.

Analysis of data indicated no significant increase in skills transfer for students in the treatment groups who used the home remodeling simulation with contextual advisement ($r = .111$, $p = .325$). Skills transfer as measured by pre-test and post-test scores for the four treatment groups and the control group were as follows.

- Control group, no simulation, contextual advisement or competition, 41.
- Treatment group 1, contextual advisement with competition, .47.
Treatment group 2, contextual advisement without competition, .82.

Treatment group 3, competition without contextual advisement, .78

Treatment group 4, no contextual advisement and no competition, .25.

The authors reported that a 2 x 2 ANOVA of the four treatment groups did not yield statistically significant results. Therefore, the authors' hypothesized outcome that contextual advisement and competition would promote skills transfer was not supported. At the same time, the authors did discover a significant interaction between competition and contextual advisement. Students in the treatment groups with contextual advisement achieved greater skills transfer when competition was not present, while students in the treatment groups with competition performed better when contextual advisement was absent. This would seem to support the authors' hypothesis that competition interferes with effective learning and may inhibit attention, elaboration and meta-cognitive skills. While competition may be appropriate in some instructional situations, i.e., educational games, it does not appear to be a factor which contributes positively to learning in educational simulations (Alessi and Trollip, 2001).

This project seems to indicate that simulations can facilitate learning transfer in a classroom setting in which the simulation approximates a situated learning context. By removing the element of competition present in many traditional classroom activities and video games, students were able to approximate a "community of practice" experience (Hung and Chen, 2001). The work of Van Eck and Dempsey (2002) lends support to the efficacy of simulations as a vehicle for delivery of situated learning experiences to the traditional classroom. Confirmatory studies discussed by Cole et al., (2006) have established the benefits to the learner of a media-rich, interactive,
computer-based simulation environment above paper and pencil simulations. In addition, these studies have demonstrated that in some instances, students learn more effectively when they are called upon to make decisions in an information-rich environment. One-dimensional pencil and paper simulations cannot provide this type of learning environment. Finally, computer-based simulations have been shown to be more adaptable than pencil and paper simulations to individual skill levels and learning styles.

*Teaching Probability with Simulations*

This study reports on the Probability Inquiry Environment (PIE) simulation, an effort to teach probabilistic reasoning through student-controlled collaborative inquiry activities (Russell and French, 2002; Vahey et al., 2000). The project goal was to encourage student construction of concepts through interaction with a student-controlled simulation. The simulation was designed to guide middle school students toward a deeper understanding of probability theory. Probability is integral to many academic disciplines i.e. biology, genetics, and psychology.

The treatment group was two seventh grade classes ($n = 45$) in an urban middle school who used the PIE curriculum for three weeks. Each of the activities during the three week curriculum included a simulation-based and hands-on collaborative learning exercise. Students worked in pairs and subsequently participated in a class-wide reflective discussion. A control group of two classes taught by the same teacher as the PIE group ($n = 54$) covered the same topics as the treatment group but used traditional curriculum rather than simulation and hands-on activities. Both control and treatment groups receive the same post-test at the end of the three week period.
Data analysis included a 3-way ANOVA carried out on three between-subject factors on the posttest: condition (experimental or control), gender and test score. Statistically significant main effects were revealed between the treatment and control groups ($F = 9.7, p < .01$), and for standardized test level ($F = 45.7, p < .01$). The main effect of gender was statistically non-significant ($F = 1.3, p = .25$), and no interaction effects were evident. Additionally, t-tests comparing both groups on the pretest were statistically non-significant ($t(89) = .21, p > .5$). However, a statistically significant difference on the post-test was observed ($t(97) = 3.4, p < .01$). These results support the author’s hypothesis that the PIE curriculum provided a positive learning experience for students regardless of academic achievement or gender. Further, statistical results indicate that students who used the PIE curriculum performed better than students in the control group. The authors asserted that simulations coupled with hands-on activities can help students to articulate difficult concepts.

The authors conclude by stressing the importance of the human teacher or tutor when using simulations and hands-on activities in a collaborative environment. The human teacher can help guide learner discussions towards critical thinking, hypothesis testing and articulation of critical results.

**Astronomy and Geology Simulations for Middle School Earth Science**

The effectiveness of simulations in science education is a topic of continuing debate in the literature. Jackson (1997) documented three case studies focusing on use of simulations in middle school astronomy and geology instruction. The integration of interactive video into these simulations was a key factor in the author's selection of the three case studies reviewed. "Microcomputer and interactive video simulations can
facilitate student learning in science by providing feedback about natural or technological phenomena which can be used to test hypotheses or models for scientific exploration and understanding” (Jackson, 1997). The simulations used in this study were all designed to address preliminary earth-science and astronomy topics appropriate to middle school science class.

The setting for this study was a public middle school located in a small urban environment in the south eastern United States. At the time the study begin the middle school possessed extremely limited computing resources. All equipment and software used in the study were provided to the school through seed funding from the National Science Foundation sponsored Statewide Systemic Initiative (SSI) project (see http://www.nsf.gov/). Three science classes were selected for the study, ranging in size from 15 to 35 students with median and average class sizes of 28 and 27 respectively. Sample demographics, based on observation, indicated approximately equal distributions of males and females in all classes. The percentage of African American students, both males and females, ranged from 40 to 90 percent. Of the three instructors who participated in the study, one was a white male and two were white females. The author was present and observed all class sessions discussed in the study, covering a two-day period. Although more than one student used the simulation software in each science class, the author chose to treat each class as a single case. Preliminary data for the study was collected from semi-structured interviews with the three teachers involved, student records, videotape recordings of typical class activities focusing on teacher-student interaction, and field notes taken by the author. Teachers were interviewed again after observing students using the simulations.
From his field notes the author reproduced a number of informative comments from the instructors involved in the study. First, instructors noted that students demonstrated difficulty navigating the command and menu structures of the simulations. One instructor recommended that students be given adequate time to become familiar with the mechanism of the simulation software before attempting effective learning activities. Instructors also commented that students became easily lost in the simulation environment without timely guidance and supervision. All three instructors observed positive motivation aspects of the simulations in that most students became very enthusiastic about the material being taught. However, one instructor stated that, due to the technical complexity of the simulation, she was not comfortable using them as anything more than a supplementary learning tool.

The author drew two key conclusions supported by data collected during the study. First, instructional simulations can motivate students to become actively involved in their own learning. Second, the positive instructional benefits of simulations can only be realized if students are guided and supervised throughout their exploration of the simulation. The presence of an instructor or tutor is crucial for effective use of simulations in instructional settings.

This study would seem to reinforce two key aspects of instructional simulations which are already well documented in the literature. First, these technologies can motivate students to become more actively involved in their own learning. Second, without proper guidance in the form of a teacher or tutor, students may undergo a negative learning experience when using these technologies. These results are supported by Lazarowitz and Huppert (Spring 1993) in a similar study among high
school biology students. Multi-user virtual environments (MUVEs) have been demonstrated to be viable for providing laboratory simulations as discussed by Clarke and Dede, (April, 2005) and Ketelhut et al., (April, 2006). The next project presented continues the present discussion of simulations for teaching middle-school earth science.

River Ecosystem Simulation for Middle School Earth Science

Dwyer and Lopez (2001) reported observations of elementary and middle school science students engaged in exploration of a river ecosystem using a simulation with integrated interactive video. The interactive video was used to realistically depict various aspects of the river ecosystem, including organisms and life cycle processes. The authors evaluated the role of the simulation throughout all aspects of the learning cycle.

Participants in the study were 31 students selected from a private middle school science class, all of whom exhibited some form of documented learning disability. Most students used their own laptop computers to run the simulation software while those students who did not own computers used desktop computers provided by the school. This arrangement resulted in a 1:1 student to computer ratio. Data for the study was collected from three surveys developed and administered by the science teacher as well as comments from her daily journal. These surveys primarily measured student attitudes towards using computers. In addition, students were videotaped while using the simulation. Additional data were collected from interviews with students and the teacher, student field logs, student concept maps and outcomes from lesson activities.

Before using the simulation, all students were taught a lesson covering the basic science concepts presented in the simulation. To allow students to gain a degree of
comfort with the simulation software, the first assignment consisted of a scavenger hunt in which students were asked to identify as many organisms as possible in the simulated river ecosystem. Students were allowed to work in small groups. For the next assignment, students were asked to construct food chain models for specific organisms from the simulated river environment. Subsequent lessons became increasingly challenging requiring students to build and elaborate their knowledge base.

Based on observational and interview data collected during and after the study, the authors’ stated hypothesis that simulation can provide “a meaningful learning experience for both the teacher and the students” (Dwyer and Lopez, 2001) seems to be supported. The teacher commented that the simulation allowed students to experience realistic problems involving many variables and arrive at realistic solutions in a relatively short time period. She also noted that, while using the simulation, students were motivated to build their own knowledge models and use them to solve problems. The authors concluded that, while simulations cannot and should not replace hands-on learning experiences, they can significantly facilitate the learning of complex concepts from science and other technical domains. Soderberg and Price (2003) reported positive results from a similar study in which instructional simulations were used with university students in biology instruction. In a further study documented by Hounshell and Hill (1989), university students who used a simulation in a biology course exhibited significantly high achievement and demonstrated greater motivation than students who used a traditional curriculum.

Moving from the middle-school earth science class to the university medical school, the next project discussed examines a medical simulation. The medical course
in which this simulation was used is part of an ongoing distance education program conducted through Massachusetts General Hospital. In this case, interactive video was included in the simulation to facilitate a sense of realism and to allow large numbers of students to interact with sophisticated medical equipment.

**Medical Training Simulations**

Simulations can be incorporated into a wide variety of instructional situations. The medical education community has investigated the use of simulations incorporating video teleconferencing and interactive video to supplement traditional face-to-face instruction (Heinich et al., 1999; Jacobs and Rodgers, 1997). These efforts have been motivated by the constantly expanding curriculum of most medical training programs. At the same time, educators have investigated ways of getting medical students more actively involved in their own learning (Levison and Straumanis, 2002).

Cooper et al., (2000) reported on a realistic medical simulation project conducted by the Center for Medical Simulations, Boston, MA (see http://www.harvardmedsim.org/CMS%20Front.html). The first phase of the project was carried out at Massachusetts General Hospital on May 22, 1997 (see http://www.mgh.harvard.edu/). The project consisted of several two-way, interactive seminars in which medical cases were presented to large audiences at widely dispersed locations. Although the primary information delivery medium in this project was video teleconferencing, simulation was used for medical telemetry.

According to the authors, these medical simulations were focused on the goals of allowing students to see the effects of their actions in real-time, to enhance learning by facilitating concurrent presentation, discussion and participation at a distance. The
simulations made it possible for students to conduct hypotheses testing in real-time and
discover cause-and-effect relationships which more traditional instructional methods
might have rendered less apparent. An added benefit was that students were able to
“observe and interact with medical equipment which was in limited supply or
inaccessible for viewing by large groups” (Forinash and Wisman, 2001). In this case,
interactive video added a degree of “immediacy and authenticity” to the simulations that
would have been difficult to achieve otherwise (Aldrich, 2001).

Although the authors did not conduct a large-scale assessment of the simulation
project, a survey instrument was administered to one of the largest audience groups to
assess user reactions to the methodology. Survey responses were generally
enthusiastic with regard to the technology, although some respondents questioned the
cost-benefit ratio. In a very similar study reported by Lane et al. (September 2001),
results indicated that use of simulations, “offers an alternative to learning with real
patients and allows a wide range of skills to be practiced and mastered.” In another
similar study involving veterinary students Farnsworth (1996) observed that repeated
use of the simulations resulted in increased students’ diagnostic efficiency. Harrison and
Gaba, (June 2005) documented yet another study conducted at Stanford affiliated VA
Palo Alto Health Care System in Palo Alto, California. The simulation instructed medical
interns in patient-care safety. Most students reported having a very positive learning
experience with the simulation.

Further experimentation with this type of simulation has been held back due to
the high bandwidth requirements of video teleconferencing. However, the development
of streaming video may facilitate future research projects of this type (Gralla, 1999).
Cooper et al., (2000) stressed the need for development of additional simulations so that more comprehensive data could be collected as to the methodology’s efficacy.

Medical simulations are perhaps some of the most sophisticated application of simulations today. However, no matter how complex the simulation, its effectiveness will be significantly curtailed if the simulation does not facilitate appropriate and timely feedback for the learner. The importance of feedback in instruction has been well documented in the literature. The role of feedback in instructional simulations is documented in the follow two studies, physics instruction and software engineering simulation.

Simulations for Physics Instruction

Zacharia and Anderson (2003) reported on using a simulation to instruct pre-service and in-service teachers in methods for teaching physics. Participants in the study were 13 postgraduates registered for a survey course in physics. Four of the participants were in-service science teachers and nine were pre-service science teachers. The sample demographic consisted of five males and eight females ranging in age from 24 through 47. None of the participants were physics majors, had taught physics or had taken a physics course within the two years preceding the study. All of the 13 participants had at least one year of introductory college physics, but none of them had taken a physics class within two years prior to this study. Although the participants were computer literate, an orientation to the simulation was given in which problems were presented similar to those which would be encountered in the simulation scenario.

The authors claimed that small sample size (n = 13) necessitated the use of a
self-control design. Participants were randomly assigned to either the experimental group (S) or the control group (N). Group S used the simulation while group N was assigned additional problems from the course text. A sequence of 12 topics were covered in the course. For each topic, the two groups were alternatively assigned the simulation or textbook problems. This resulted in 13 participants multiplied by 12 topics for a total of 156 cases, 78 cases in which each student used the simulation and 78 cases where each student used the textbook-based curriculum. Participants were asked not to discuss the specifics of the learning experience among themselves. In addition, participants were asked not to use instructional materials other than those presented in class and in the course text.

The authors of this study concluded that the simulations contributed to a positive learning experience for the participants. Participants demonstrated improvement in their ability to predict scientific phenomenological outcomes as well as render scientifically correct explanations of experimental phenomena. Further support for this conclusion can be found in Thomas and Milligan, (2004). The simulation created an environment in which participants’ preconceptions about physics concepts could be challenged.

“Computer simulations are an excellent tool for fostering conceptual change “ (Zacharia and Anderson, 2003). The authors concluded with a call for additional research focusing on how maturity and prior experience impact perceptual change when interacting with simulations.

Software Engineering Simulation

Sharp and Hall (2000) reported on a case study of a software engineering course offered through the Open University in the United Kingdom (UK). Students enrolled in
the course interacted with an instructional simulation of a software publishing house. The object of the simulation was to give students a situated learning experience by participating in a software development team in a real-world setting. The simulation allowed students to make choices based on incomplete information and to see the consequences of their choices. The learning objective of the simulation was anchored instruction, resulting in a self-motivated and relevant learning experience (Heinich et al., 1999). Anchored instruction has been defined as situating the learner in a realistic problem setting where the learner can experience the same types of problem scenarios a domain expert would encounter (Cognition and Technology Group at Vanderbilt, 1990, 1992; Cohen et al., 2006). Problem scenarios incorporate real data, are factually authentic and present the learner with realistic tasks which are likely to be confronted in a real-world situation. Anchored instruction is similar to both case-based and problem-based learning (PBL). However, anchored instruction differs from case-based learning in that problem scenarios are designed to be discussed and experienced rather than passively observed or studied. PBL differs from anchored instruction in that learners are confronted with conducting their own research outside the PBL scenario prior to confronting the problem. Anchored instruction scenarios usually embed all data and resources needed by the student to address and resolve the problem. In this regard, anchored instructions is a less open-ended learning method than PBL. The materials presented in the software publishing house simulation were supported and supplemented by a course pack of printed materials sent to each student. The course pack contained basic information while the simulation provided challenging applications.
of basic knowledge and skills. In addition, each student was assigned to a tutor/coach who provided scaffolding by email or telephone.

Feedback is an essential element in any learning experience (Gagne et al., 1992). In the case of the software publishing house simulation, a significant feedback element was provided by the simulation software itself. This feedback was presented to the student in the form of suggested solutions to sample problems.

To gauge the effectiveness of the simulation, the authors collected user feedback through questionnaires and usability studies. Although results were mixed, in general students viewed the simulation as engaging and easy to use. Positive responses to the simulation focused on the inclusion of real-world case studies. Negative responses addressed the relevance of the simulation’s user interface to the course and the amount of time required to work through the course pack.

Data collected by the authors supported their hypothesis that students recognized and valued the real-world learning experience facilitated by the simulation. Indeed, in a similar study reported by Kumar (1996), learning outcomes for students using an instructional simulation were almost identical to students engaged in hands-on assessments. Tang (2002) documented a study in which business students interacted collaboratively with a simulation of a small manufacturing plant. Variables examined in the study were effectiveness of training protocol, feedback and skill retention. Statistical results reported by the author seem to indicate that collaboration facilitates skill retention. Skills retention measured immediately after completion of training yielded non-significant results \( F(1, 56) = 2.63, p > 0.01 \). Skills retention seemed to increase after one week \( F(1, 56) = 3.73, p > 0.05 \) becoming statistically significant after one
month \((F(1, 56) = 9.55, p < 0.01)\). These data support the author’s hypothesis that students will benefit from engaging with instructional simulations in a collaborative learning environment. Whitecotton (2003) documented a similar study in which 223 students enrolled in an undergraduate Computer Science course at Sam Houston State University used an instructional simulation in place of the traditional text-based curriculum. Results of the study indicated a significant increase in the number of students who received A's, a distinct decrease in the number of students who received B's, C's, or D's, and a moderate decrease in the numbers of students who received F's.

The next 4 studies presented discuss the facility of discovery learning which instructional simulations can provide. These studies also consider the deployment of instructional simulations over the Internet.

**VR-enhanced Simulation for Computer Graphics Instruction**

Some researchers and educational practitioners have explored the use of virtual reality (VR), immersive virtual environments (IVE or VE) or virtual worlds (VW) as means of facilitating constructivist learning activities (Blascovich and Bailenson in Cohen et al., 2006; Briggs and May, 2002; Miettinen, 2002). Increasingly, simulation designs incorporate IVE interfaces to facilitate participant learning and interaction. IVE's can contribute to the participant's phenomenological experience of immersion in the simulation scenario. The capabilities IVE’s provide for exploration and manipulation of objects and procedures as part of a simulation have significant implications for learning with simulations. IVE’s can also present the simulation participant with a volume and fidelity of sensory data very closely approximating a situated learning experience.

Sung and Ou (2001) reported on an Internet-based computer graphics course in
which VR technology was incorporated into a simulation with the goal of increasing learning effectiveness. The computer graphics course discussed was offered through the Department of Electrical Engineering, National Central University, (see http://www.ncu.edu.tw/English/) Chung-Li, Taiwan. The authors asserted that the learning experience provided by a VR-enhanced simulation was more meaningful because it was derived from the student’s own exploration of the simulation environment (McLellan, 1996).

Preliminary analysis of the effectiveness of the VR-enhanced simulation was determined by administering a pretest and posttest both to students who had used the simulation and students who had used more traditional learning systems. Test results indicated that students who used the VR-enhanced simulation scored higher on practical examinations (posttest). The authors also observed that students who had access to the VR-enhanced simulation returned frequently to the course Web site to refresh their skills. In addition, these students were observed to retain a higher level of cognitive knowledge than students who had not used the simulation. Students reported that using the VR-enhanced simulation was a rewarding and positive experience. Further, students viewed the simulation more as a gaming activity with instructional value than as a homework assignment. As a result, students were willing to devote more time working in the simulation than they would have allocated to more traditional study.

The authors concluded by stating that VR-enhanced simulations provided students with “experiential learning” (Sung and Ou, 2001). This type of high-level learning was possible because the simulation involved the student in active completion
of specific tasks and complex operations. Research has shown that students involved in “experiential learning” tend to remember 90% of what they encounter in the course of the learning activity (Heinich et al., 1999). If VR-enhanced simulations can provide this type of learning experience over the Internet, then they may have a place in the classroom as well (Ryhme, 2002). However, it is crucial that students not view an instructional simulation activity merely as a gaming activity. Interacting with a simulation as if it were a game will probably result in a less than optimal learning experience.

**Java Applet-based Microworlds**

A variety of technologies may be used to deliver instructional simulations through the Internet. Min (2001) reported on an ambitious project conducted at the University of Twente in the Netherlands (see http://www.utwente.nl/en/) in which Java applets were used to deploy simulations to students in an array of distance settings. These Java applet-based simulations were designed to be downloaded over the Internet and were referred to as, “microworlds” (Min, 2001).

The simulations developed by Min and colleagues were strongly grounded in constructivist learning theory and Vygotsky’s zone of proximal development (ZPD) concept (Litowitz, 1993; Miettinen, 2002; Sternberg, 1999). Min stated that these were model-driven stand-alone simulations as distinguished from drill and practice or tutorial courseware. The simulations were designed to fulfill the rolls of in-class demonstration, coached learning, individual discovery learning, interactive practice, and assessment. Of equal importance, the simulations provided a vehicle for evaluating whether students successfully applied what they had learned to practical problems. Min also stressed the vital role the simulations played as a feedback channel for the instructor.
Min stated that successful application of simulations demanded “coached learning, two-way communication, feedback, demonstrated ability of the student to form sound hypotheses, access to appropriate manuals, written assignments and well-designed printed materials”. Without these elements, the simulations could not have achieved the author’s instructional goals. Min observed that the more time students spent working in the simulations the more the students learned. However, Min also noted that if the student used the simulations without coaching, the result was often ineffective practice. Further, if the student interacted with the simulations without first mastering the appropriate problem related skills, the interactive practice often resulted in null or incorrect learning.

The simulations developed by Min and colleagues supported discovery learning through the use of cases and scientific experiments. Each case was presented to the student in print format with the essential dynamic elements of the case portrayed by the simulations. The student was then directed to construct and test a hypothesis, and manipulate one or more parameters of the simulation until the simulation model behaved normally. Scientific experiments were designed to resemble a vocational practicum in which the student measured specific variables with the goal of attaining a certain insight. Results of the experiment were recorded by the student in an electronic or paper worksheet. The student then constructed charts or used other graphic representations to visualize the results of the experiments. Finally, the student evaluated the resulting visual representations with the goal of gaining in-depth understanding of a real-world phenomenon.

While Min did not report quantitative data on use of Java applet-based
simulations, the potential of the technology is clear. Given the relatively small file sizes of Java applets and the ease with which they can be accessed through the Internet, the Java applet simulations described by Min may be applicable to a wide range of classroom and distance learning situations.

Simulations for Distance Learning

The Internet may prove to be an optimally functional vehicle for delivering instructional simulations to students at a distance (Simonson et al., 2003). Granland, Bergland and Eriksson (2000) reported on the development of three Internet-based simulations for the Department of Computer and Information Science, Linköping University, Linköping, Sweden. The authors focused on the relationships between Internet-based simulations and instructional strategies appropriate to simulation-based learning environments. A number of learning methodologies for which simulations may provide optimal learning outcomes were presented. These methodologies included problem-solving, demonstration, experimentation, exploration and hypotheses testing.

The three simulations presented by Granland, Bergland and Eriksson were named, Chernobyl, C³ Fire and ERCIS. All three simulations were implemented using Java applets. The Chernobyl simulation was designed to teach basic operations of a nuclear power plant as well as rule-based modeling. The simulation introduced plant operations and allowed the student to deal with certain malfunctions which can occur during the course of normal plant operation. The Chernobyl simulation included three prewritten cases and one random case in which the course of events was not determined in advance. While the simplified physics model on which the simulation was built was inaccurate, the simulation did familiarize the student with the dynamics of a
The C\(^3\) Fire simulation was designed to present “Command, Control and Communication” problems in an Internet-based learning environment. The goal of the simulation was to let the student experiment with various strategies for team training, team coordination and situation-awareness. The metaphor for the simulation was fighting forest fires and included fire fighting units, vegetation, houses and other simulated agents. Fire-fighting was used merely as a vehicle to demonstrate the problem solving principles inherent in team management. C\(^3\) Fire was designed to allow the learner to experience some of the dynamics present in a real-world emergency situation. To facilitate this, the fire-fighting scenario played out by the simulation changed autonomously and in response to the learner’s actions. The simulation maintained a detailed log of the learner’s actions and reactions to the changing scenario. This log was later used by the instructor to evaluate the learner’s performance.

A project quite similar to C\(^3\) Fire was presented by O’Looney and Dodd in Cohen, et al, (2006). The Disaster Control simulation project was developed by the Carl Winston Institute of Government at the University of Georgia in cooperation with the Georgia Emergency Management Agency. The goal of the Disaster Control simulation is to provide students with the experience of a realistic disaster management situation within a safe environment. The simulation allows students to create and test problem-solving skills to manage a rapidly changing crises situations in the face of time constraints, limited resources, information overload and problematic communications.

ERCIS and C\(^3\) Fire both utilized a team learning environment as opposed to a
single user simulation demonstrated in Chernobyl. ERCIS simulated certain key aspects of the RBS-70 unit of the Swedish Anti-Aircraft Defense. The goal of ERCIS was to provide “training” with equipment and procedures related to the RBS-70 unit (Noble, 2002). The simulation abstracted some aspects of the RBS-70 technology, focusing on key functions and operational parameters relevant to group activity in a real-world combat setting.

All three simulations presented by Granland, Bergland and Eriksson shared the goal of helping the learner distinguish between conceptual and operational knowledge. Subsequent to observation and evaluation of the simulations, the authors concluded that the Chernobyl simulation provided the learner with a good understanding of the underlying model upon which the simulation is built. The $C^3$ Fire simulation allowed the student to learn about various aspects of a dynamic situation where the model underlying the simulation was not the focus of instruction. The ERCIS simulation facilitated mastery of operational knowledge in a situation where the concepts and user interface of the simulation model closely resembled a real-world setting. All three simulations focused on discovery learning in which the student explored the simulation environment, collected data, analyzed information, and made informed decisions in order to acquire knowledge. Further, the simulations were designed to emphasize affective learning, incorporating as much motivation appeal as possible.

Although all three simulations were designed for use in a distance setting, the authors stressed the need for teacher-guided learning and appropriate feedback (Notar, Wilson and Ross, 2002). While the authors presented no data in support of their research, they asserted that Internet-based simulations have two key advantages. The
first and most obvious advantage is that simulations built with Java applets are easily and widely accessible to any student with Internet access. Java applets can provide the flexibility to “address different learning styles and provide access to a variety of media elements” (Roccetti and Salomoni, 2001). Second, and more importantly, simulations can present the learner with opportunities to experience dynamic and interactive environments. The value of “experiential learning” has been well documented in the literature (Heinich et al., 1999). If well-designed, model-based simulations can be made available over the Internet, students in any setting can engage in “real-world problem-based learning” (Notar et al., 2002).

Microelectronics Instruction Simulation

Hensgens, van Rosmalen and Hanu (1998) reported on the MODEM project (Multimedia Optimisation [sic] and Demonstration for Education in MicroElectronics) (see http://www.ecotec.com/sharedtetriss/projects/files/modem.html), an effort to support active learning through the use of simulations. MODEM was developed at the Research Institute for Knowledge Systems (RIKS bv), Maastricht, The Netherlands. The goal of the MODEM project was to allow students to acquire complex knowledge and skills relevant to the microelectronics industry through experience with professional microelectronics modeling software tools. Through hands-on experience with real-world tools in a simulated work environment, students were able to explore and experience the key concepts of microelectronics modeling.

The MODEM simulation incorporated access to real-world software tools which ran on a server. This was facilitated by a software bridge which connected desktop PCs to a UNIX server using the PC-X-server, HCLEXceed. Interactive multimedia and
hypermedia were extensively utilized throughout the MODEM simulation to support constructivist learning, interactivity and maximum learner control. The simulation promoted learning by doing; students were free to make mistakes and acquire knowledge from solutions they developed. Further, the MODEM simulation motivated students to build and test their own hypothesis acquiring high-level knowledge through development of complex problem-solving skills. Finally, because the MODEM simulation was delivered over the Internet, it eliminated the same-time same-place constraint present in more traditional microelectronics instruction.

The MODEM simulation software was developed using Asymetrix ToolBook® authoring tool (see http://www.asymetrix.com/en/toolbook/index.asp) and designed to run in either Netscape® or Internet Explorer® Web browsers. Synchronous communication among students and between students and teachers were facilitated with Microsoft® NetMeeting (Microsoft Corporation, www.microsoft.com). The whiteboard functions of NetMeeting were used extensively for feedback and collaborative work. While the MODEM simulation was designed to provide the learner with full control over the course materials, extensive feedback, scaffolding and coaching were provided through NetMeeting. The authors stressed that coaching and scaffolding were essential to prevent the learner from becoming lost in the simulation.

Evaluation of the MODEM simulation was conducted at the University of Leeds and the University of Twente. Data was collected by administering usability questionnaires to a small group of software testers. The authors reported an overall positive response to the simulation, although no quantitative data were provided. Subsequent testing was carried out with students, all of whom were experienced
computer users. Upon completion of the microelectronics course, the students were asked to complete the same usability questionnaires previously presented to the software testers. Once again, the authors reported very positive student response to the simulation, but provided no supporting data. In particular, students commented most favorably about instructor feedback made possible by the NetMeeting software. Although some technical communication problems did arise during the course, students worked around these difficulties and did not consider them a negative aspect of their experience with the MODEM simulation.

The authors emphasized that the MODEM simulation was unique because it incorporated access to real-world resources and was built partially around existing software, i.e. NetMeeting. That part of the simulation constructed with ToolBook was designed to bring the preexisting software packages together under a coherent user interface, provide consistent and relevant feedback, and give students complete access to all course materials. The ToolBook user interface also facilitated note taking and collaborative work. The authors stated that MODEM represented a viable and cost effective approach to the development of simulations for the classroom as well as blended or distance learning venues. Similar results were reported by Jurica et al. (1999) from a study in which two instructional simulations were integrated into a blended learning setting.

*Environmental Science Instructional Simulation*

Kelsey (2002) documented a simulation project titled, Brownfield Action which was developed at Columbia Center for New Media Teaching and Learning (CCNMTL) from 1999-2001. The simulation was deployed and tested at Barnard College, a
women’s college affiliate of Columbia University. The project goal was to demonstrate how a simulation could be used to integrate laboratory and lecture components in an undergraduate Environmental Science course. The simulation was used throughout the semester and presented students with complex problems in Environmental Science. Addressing these problem tasks required the students to apply skills and knowledge acquired in lecture and laboratory setting. To successfully complete the semester-long simulation, the students were required to become actively involved in both laboratory and lecture activities. With each new problem encountered, students had to form theories, hypothesize, engage in discovery activities and apply acquired knowledge to arrive at a valid conclusion.

The Brownfield Action simulation was designed as a learning environment in which students could learn to apply key concepts from environmental science. Students operated a mathematically driven model of a town, taking on the role of a consultant to a real estate developer who seeks to purchase land in the town. The key component in the town model was a contamination event. Accordingly, the real estate developer wished to avoid purchasing potentially contaminated land. Students had to become fully engaged with the role and associated responsibilities to gain maximum benefit from the simulation. In so doing, students could not only compile a set of skills transferable to real-world science activities, they also gained an appreciation for the connection between science and real-world events.

At the beginning of the semester, students were surveyed to obtain a measure of their perceptions of course topic as well as to gage their existing knowledge relevant to the course. At the end of the semester, students rated the simulation on how effectively
it contributed to course objectives. Subjective data were collected from hour-long focus groups in which students were asked to reflect upon their interactions with the simulation and to articulate their learning experience.

While no statistical data were provided, the author concluded that students who used the Brownfield Action simulation learned more than in previous semesters. Further, assignments completed by students more closely resembled results expected from a science professional. The simulation appeared to have increased scientific literacy, facilitated and expanded student's theory construction skills, and increased student learning overall.

The author reported two notable problems with the Brownfield Action simulation. First, inadequate staff development time was built into the simulation deployment schedule. Instructors did not have time to become sufficiently comfortable with the technology. As a result, they were not able to help students engage in discovery learning activities to the fullest extent. Second, student frustration over technical glitches in the simulation prevented complete realization of some learning objectives.

In concluding comments about Brownfield Action, Kelsey stated that, “the software itself does not teach . . . it invites the user to engage in the problem”. Successful engagement requires motivation and a willingness to take risks in a safe environment like the classroom. Cohen et al. (2006) have documented the fundamental role of motivation in any successful learning experience. Brownfield Action exemplifies how a motivating and engaging learning experience can be achieved.

*Integrating Simulations Into Homework Assignments*

Ronen and Eliahu (1999) reported on a pilot study which focused on integrating
simulation-based activities into existing homework assignments. The setting for the study was an urban middle-class neighborhood school. Approximately 75% of the student population owned personal computers. The sample size was 71, all 9th grade students from two different classes, taught by different teachers. The sample was composed of 30 boys and 41 girls. All participants in the sample had completed a mandatory course in basic computer literacy. Over a period of two months, students engaged in simulation-based homework activities as part of their study of science. Students’ reactions to the simulation activities were analyzed and correlated to their prior academic performance and computer experience.

The simulation software used in this study was designed by the authors and titled, *DC-Kit* (Ronen and Eliahu, 1997). The simulation was specifically designed to provide learning activities on electric circuits, focusing on students’ common misconceptions. All activities were structured to provide constructive feedback with the goal of helping students evaluate their own work. The user interface was crafted to resemble a simplified version of what one would find in most computer aided design (CAD) programs. Both qualitative and quantitative problem-solving activities were included in the simulation. Weekly simulation-based homework activities were similar to normal homework assignments in regard to time and skills requirements. During the two month study the simulation was used only once in class. In this instance, the simulation was used as part of an experiment in a physics lab.

Data were collected from a questionnaire, interviews with students, interviews with teachers, and a final exam. The authors reported that 47% of students completed all simulation-based homework activities, 36% finished most tasks and 17% completed
only some activities. Based on self-reported data, most students spent 1-5 hours working with the simulation each week. Fewer than 10% of students worked with the simulation more than 5 hours per week. Statistical analysis revealed a high degree of correlation between time, the number of completed simulation-based homework activities and exam scores ($R = 0.3$, $p < 0.05$). Based on data collected from interviews approximately 50% of participants stated that simulation-based homework activities motivated them to spend more time on assignments than they would have otherwise. These findings are supported by very similar studies documented by Hounshell and Hill (1989) and Shaw et al., (1985). Authors of these studies reported increased motivation and content engagement, improved problem-solving, logical reasoning skills, knowledge transfer and emergence of metacognitive and knowledge seeking skills on the part of students who used instructional simulations as part of their science laboratory assignments or homework exercises. Results of cognitive research studies indicate a synergistic relationship between engagement and successful learning (Quinn and Conner, 2005). Additional confirmatory studies have been produced by Bourque et al., (March, 1987), Helgeson, (1988), Rivers and Vockell (May, 1987).

An Assessment Instrument for Instructional Simulations

A number of research projects have thus far been presented focusing on the use of instructional simulations in the classroom as well as in an open or distance-learning venues. Although the authors of these projects have attested to their success, the claims of some authors have not been supported with quantitative data. The collection of quantitative data is made problematic by a lack of an assessment paradigm for instructional simulations (Milrad et al., 2000). The need for assessment instruments with
documented validity and reliability is crucial (Feinstein and Cannon, 2002).

Dean and Webster (2000) examined an interactive simulation in the context of a distance education business degree course. Their goals were to develop an instrument to assess whether simulations motivate “high quality learning”, and to determine whether simulations impact students’ ability to transfer knowledge to the real-world. High quality learning is essential for moving the student to a state of “metacognition” where the student takes responsibility for his/her own learning (Sternberg, 1999). The authors asserted that the variable and inconclusive results obtained with existing assessment instruments pointed to the need for new assessment tools geared toward simulations in distance settings. Because simulations tend to focus on student-centered learning, the authors stated that any new assessment instrument needed to be more focused on student-related factors.

The simulation used in this study was designed to support development of cognitive models, provide interactive practice, encourage hypothesis formation, hypothesis testing, experimentation and mastery of concepts through application of knowledge to real-world problems. The simulation involved the student in theory-and-practice exercises with the goal of enabling the student to apply acquired knowledge to realistic work environments. The simulation software was built on a decision support system and tutorial which encouraged the student to apply acquired knowledge to work-based decision making. Interactive practice was achieved by allowing students to make their own decisions through a series of scenarios presented by the simulation. Students received feedback about their decisions and guidance with regard to factors not considered. Direct face-to-face interaction with instructors or other students was very
limited during the study. Most feedback came to the student through the simulation.

The authors conducted an assessment by distributing a survey instrument to 150 students who had completed the business course using the simulation. Detailed quantitative results presented by the authors indicated that current simulations do not promote transfer of knowledge to a greater degree than other methodologies. In short, simulations as currently constructed for distance education do not appear to facilitate transfer learning of acquired knowledge to real-world situations. At the same time, survey results indicated that students responded positively to the high degree of interactivity. In this regard, simulations do appear to have a positive impact on students’ motivation to study.

The authors encouraged others to develop similar instruments for assessment of simulations delivered to students in the classroom and at a distance. However, the authors cautioned that such instruments should be carefully crafted to focus on student-related factors as well as factors pertaining to cognition, transfer learning and motivation. As an alternative to creating a new assessment instrument, Hertel and Millis (2002) have suggested using peer reviews of simulations, participant questionnaires and self-evaluations, as well as focus groups.
CHAPTER 3

METHODOLOGY

Need for Replication

The use of simulations for classroom instruction can be traced to the 1960s (Taylor, 1987). However, research on the effectiveness of simulations in the classroom as aids to instruction remains limited and to some degree problematic. Much extant research suffers from poor research designs, inadequate sample size, inappropriate data analysis techniques or lack of informative data reporting (Akpan, 2001). Further, the attention of simulation designers and practitioners seems focused more on the development of new simulations than meaningful assessment of those already in use (de Jong and van Joolingen, 1998). The problem has been exacerbated by the absence of a standardized design model for instructional simulations as documented in Adams, Reid, LeMaster, McKagan, Perkins and Wieman, (2007a, 2007b). Accordingly, further studies are warranted in order to gain a more complete understanding of how students are impacted by their interactions with simulations as well as the role simulations can play in the curriculum.

Data collected by the National Center for Education Statistics (NCES) in 2003 indicate that less than 17.8% of university students in the United States completed degree programs in any branch of science, mathematics, engineering or technology (SMET). This number is lower than for most other developed nations, for example, Finland, 28.0%, Germany 30.1%, South Korea 37.8%, United Kingdom 31.2%) (NCES, 2005). Furthermore, the percentage of students in the United States who completed degree programs in SMET has declined from a high of 21.7% in 1985 (NCES, 2005).
According to data published by the National Science Foundation (NSF) in 2006, the number of undergraduate and graduate degrees in SMET fields has increased since 1995 to an average of 32% (NSF, 2006). However, these same NSF data indicated that as of 2002 almost 70% of Americans were unable to explain what it means to study something scientifically (NSF, 2006). These data also indicated that less than one-third of all K-12 and undergraduate students in United States public schools are able to meet their grade-level proficiencies in mathematics and science (NSF, 2006). If simulations in the classroom can stimulate interest and motivate students to explore career opportunities in SMET fields, then the expenditure of time and resources to develop such simulations seems worth the investment.

Confirmatory studies have established that instructional simulations can help students master difficult concepts from science and mathematics. See, for example, Dwyer and Lopez (2001), Jackson (1997), Kelsey (2002), Ronen and Eliahu (1999), Van Eck and Dempsey (2002) and Zacharia and Anderson (2003).

The questions addressed by this study focused on whether simulations contribute to instruction by facilitating transfer of knowledge and skills. This study did not compare instructional simulations to traditional classroom instruction as this would have constituted yet another media-comparison study (Clark, 1983). Such studies have been shown to produce little in the way of significant statistical results (Cohen et al., 2006; Nathan and Robinson, 2001).

Instrumentation

A pilot study was conducted during the fall 2006 semester on the campus of Midwestern State University (MSU) to test data collection methods and to correct procedural flaws.
(Gall, Gall and Borg, 2003; Isaac and Michael, 1997). See Appendix H. The full study was conducted at MSU during the spring 2007 term. The study focused on one section of GNSC 1204 Physical Science, a course in physics and chemistry for non-science majors. GNSC 1204 is part of the core curriculum requirements for all undergraduates in the West College of Education at MSU. The course was taught by an Associate Professor of Education and Reading and yielded a convenience sample of \( n = 78 \).

Scores were collected from four content-specific pretest-posttest scales administered by an Associate Professor of Education and Reading. A first pretest and posttest were administered early in the semester to measure student’s pre-existing knowledge of motion and force. This first pretest-posttest represented the first of three control components in the study. Shortly thereafter, a second pretest was administered focusing on sound waves. Following this pretest, students interacted with the Properties of Mechanical Waves® simulation (Knight, San Francisco, CA : Addison-Wesley, wps.aw.com/aw_knight_physics_1/0,8722,1123696-,00.html), a physical science simulation focusing on sound wave propagation. A posttest was administered after students completed the simulation. Curriculum activities centered on the sound wave simulation constituted the treatment phase of the study. Interaction with the simulation was followed by a posttest. The source for the Properties of Mechanical Waves simulation was the Web site and text, Physics for Scientists and Engineers: A Strategic Approach with Modern Physics, by Randall D. Knight, Ph.D., Department of Physics, California Polytechnic State University, San Luis Obispo, (2004). Validity and reliability of the Properties of Mechanical Waves simulation and the Knight text have been documented in the literature by Adams, Finkelstein, Gratny, Perkins and Wieman.
During the second half of the spring 2007 semester a third content-specific pretest-posttest were administered as part of the control. This assessment collected data on students knowledge of electricity. Near the close of the semester a fourth pretest-posttest was administered focusing on the periodic table. The first, third and fourth pretest-posttest scales represented the control component of the study. Students were also required to complete a lab exercise and daily quizzes through WebCT® online proprietary virtual learning environment (Blackboard Inc., Washington DC, www.webct.com). These scores were collected as documentation of attendance and as indicators of academic performance throughout the semester.

At the end of the spring 2007 term study participants completed a brief survey See Appendix E. As part of the survey students were asked to self-report on knowledge-seeking activities in which they engaged during the semester. All data were collated and analyzed using SPSS.

Research Design

A nonequivalent control group design was used for this study (Cook Campbell,
1979; Gall et al., 2003). This design is documented in Campbell and Stanley (1963) and in Isaac and Michael (1997) as the nonrandomized control-group pretest-posttest design.

\[
\begin{array}{cccc}
O_{A1} & O_{A2} & & \\
O_{B1} & X & O_{B2} & \\
O_{C1} & & O_{C2} & \\
O_{D1} & & O_{D2} & \\
\end{array}
\]

This design tested a single group of participants on four scales providing an opportunity to compare the performance of the same group of participants before and after exposure to X (Isaac and Michael, 1997). It was anticipated that one of these scales \(O_B\) would show change as a result of the treatment while the other three scales would not change. Campbell and Stanley (1963) have documented the nonequivalent control group design to be significantly more robust and less susceptible to validity threats than the one-group pretest-posttest design. The design has also been demonstrated as a suitable substitute for other pretest-posttest designs in cases where random assignment is not possible (Campbell and Stanley, 1963; Cook and Campbell, 1979). The nonequivalent control group design did not require random assignment of participants into separate control and treatment groups. Splitting the sample into two groups was not possible as a consequence of the study venue as well as limitations placed upon the study by the MSU Human Subject Review Committee (HSRC). See Appendix C. The independent variable or treatment in this study was the inclusion or non-inclusions of an instructional simulation titled, Properties of Mechanical Waves®. The treatment was presented as part of a unit on sound wave propagation midway through the spring 2007
Validity, or the truth of an inference, is a unitary concept about which evidence may be collected in support of the correctness or appropriateness of the inference (Gall et al., 2003; Shadish, Cook and Campbell, 2002). Content or logical validity was of particular salience to this study because changes in test scores served as the predominant indicator of treatment effect (Muijs, 2004). As defined by Gall et al. (2003) and Isaac and Michael (1997), content validity is a measure of the relationship between test content and the construct the test was designed to measure. Content validity must be systematically determined by content experts (Gall et al., 2003). In establishing content validity, these experts precisely define the specifics of domain content that a test is intended to assess. Once the domain is clearly delineated, the experts evaluate the accuracy with which the test samples the content domain. A content-valid test will in effect represent a definition of the content it claims to cover. For test scores to be accepted as content-valid the test must present the student with a representative sample of domain-specific content. However, it is not necessary for a single test to completely convey the entire content of a domain. Regardless of the domain, content validity can only be established by domain experts. To maximize content validity for the pilot study and the full study, the Associate Professor of Education and Reading selected test items from a test bank accompanying the course text for GNSC 1104 Life/Earth Science and GNSC 1204 Physical Science. The text used in both courses was Integrated Science, (3rd ed.), by Tillery, B. W., Enger, E. D. and Ross, F. C.. Test bank questions and domain content presented in this text have been evaluated, revised and approved for content validity by a panel of 83 domain experts from a wide range of
universities, colleges and technical institutes. The contributions of these experts are documented in the Preface to the *Integrated Science* text (Tillery et al., 2007).

Content validity was also addressed by setting the time intervals between pretest-posttest scales $O_A$, $O_B$, $O_C$ and $O_D$ to be as brief as possible. Time intervals between $O_A$ (control), $O_B$ (treatment) and $O_C$ (control) were also kept as brief as possible. Further, pretest and posttest for $O_A$, $O_B$ and $O_C$ covered topics from physics; motion and force, propagation of sound waves and electricity. $O_D$ focused on similar but non-equivalent content from chemistry, the periodic table. Finally, all pretest-posttest were similarly formatted with regard to number and type of questions and all tests had equal intervals. These measures addressed internal validity threats arising from ambiguity of direction of causality (Isaac and Michael, 1997).

**Statistical Conclusion Validity**

Validity is generally defined as evidence in support of appropriateness, accuracy or the degree to which an instrument measures what it is designed to measure (Gall et al., 2003; Sproull, 2002). Statistical conclusion validity is about errors in assessing statistical covariation (Shadish et al., 2002). The focus of statistical conclusion validity is on the degree or sensitivity of measurement of covariance among variables made possible by the study design (Cook and Campbell, 1979; Isaac and Michael, 1997). The hypothesis tested by these validity measures is whether covariance can be assumed given a specified $\alpha$ level (Cook and Campbell, 1979; Isaac and Michael, 1997). Isaac and Michael (1997) have identified 7 threats to statistical conclusion validity. To strengthen statistical conclusion validity, statistical assumptions of all analysis methods used in this study were researched and addressed where possible. Statistical
conclusion validity error rates were addressed by using standard scores whenever possible and by focusing on a single independent variable. Inferential statistical measures were not used removing the possibility of reliability of measures threats. Unreliability of treatment implementation was addressed by administering the same instruments to all participants on the same schedule. In addition, all assessments were administered to all participants by the Associate Professor of Education and Reading in the same venue. These measures reduced internal validity threats arising from interaction of selection, maturation, history and setting. Most participants interacted with the treatment in the same venue, a computer lab, reducing threats arising from random irrelevancies or extraneous variance of setting. Participants in the sample were largely homogeneous with regard to demographic factors minimizing threats to statistical conclusion validity arising from random heterogeneity of respondents.

Internal Validity

The focus of internal validity is causal reasoning errors (Shadish et al., 2002). Internal validity assesses whether or not the independent variable effected change on the dependent variable(s) as the result of a causal relationship and whether the assessment instrument in use measured what it claims to measure (Isaac and Michael, 1997; Sproull, 2002). Internal validity does not address replicability or inferences from a sample to a target population (Shadish et al., 2002). Campbell and Stanley (1963) have documented 8 types of internal validity threats while Gall et al. (2003) have documented an additional 4 threats for a total of 12. See also Shadish et al. (2002). Because most participants in this study interacted with the treatment in an open computer lab, it was not possible to prevent participants from discussing the treatment resulting in a degree
of threat to internal validity arising from diffusion or limitation of treatments. Efforts to mitigate this threat relied on providing participants with the minimum amount of information about the study while remaining in compliance with informed consent requirements imposed by MSU’s HSRC and the Institutional Review Board (IRB) of the University of North Texas (UNT). The history threat arises when a study takes place over a period of time providing an opportunity for other variables to interact with the treatment (Campbell and Stanley, 1963; Gall et al., 2003). The mortality threat arises when physical or psychological changes occur among study participants over the course of the study (Campbell and Stanley, 1963; Gall et al., 2003). The testing threat arises when improved scores on a posttest is attributable to similarities between a pretest and the posttest rather than an improvement in performance (Campbell and Stanley, 1963; Gall et al., 2003). The instrumentation threat arises when the measuring instrument has changed between pretest and posttest resulting in misleading scores (Campbell and Stanley, 1963; Gall et al., 2003). The statistical regression threat, or regression to the mean, occurs when study participants score better or closer to the mean on a posttest than on a pretest (Campbell and Stanley, 1963; Gall et al., 2003). When the statistical regression threat manifests, actual performance has not improved but merely appears to have improved as a result of scores moving closer to the mean. The most reliable defense against threats from history, maturation and mortality is to ensure that samples are as similar as possible (Gall et al., 2003). The treatment and control experienced the same history and maturation threats, encountered the same testing and instrumentation issues, and experienced similar rates of mortality and regression to the mean, minimizing these threats to internal validity (Campbell and
Stanley, 1963; Isaac and Michael, 1997). These threats were further controlled for by the presence of pretest-posttest on all four scales (Isaac and Michael, 1997). Each pretest-posttest were similar in content but with the order and type of questions changed which helped to mitigate threats arising from instrumentation and testing. The control also reduced the likelihood of misinterpreting the impact of confounding variables for the main-effect of X resulting in a Type I error (Diekhoff, 1992, p. 130-1; Sproull, 2002, p. 57). Threats from experimental mortality arise when participants drop out of the study or fail to receive part of the treatment (Gall et al., 2003). It was not possible to prevent participants from dropping the class. However, data collected from participants who missed multiple tests were omitted from the final data set providing some mitigation of experimental mortality. Finally, Internal validity threats from compensatory rivalry of the control group, compensatory equalization of treatments or demoralization of respondents arise when a sense of competition is allowed to develop between the control and treatment groups (Gall et al., 2003). This state of perceived competition can occur when study participants learn too much about instrumentation, methodology or study goals. These threats were controlled for in 3 ways. First, all participants in the study received the same treatment. Second, no participants received goods, services or other incentives to participate in the study. Third, as the control and treatment were the same group, no differential selection threat or perception of competition between groups was present.

In most studies internal validity is to a significant degree dependent upon statistical conclusion validity (Shadish et al., 2002). This relationship exists for two reasons. First, even if all statistical procedures in a study are followed with accuracy
and precision, errors in causal reasoning may still arise resulting in spurious inferences about causal relationships. The presence of statistical covariation does not in and of itself support an inference of causation. Second, statistical errors can arise in the execution of even the most meticulously executed experiment resulting in false conclusions regarding effect size and statistical significance.

**Face Validity**

Face validity is defined as the degree with which an instrument appears appropriate for its intended use, based on a visual inspection of the instrument (Gall et al., 2003; Lewis, 1999). Face validity of test items was assessed by the Associate Professor of Education and Reading by comparing test questions with specific concepts introduced in the course text and the instructional simulation.

**Construct Validity**

The focus of both construct and external validity is generalizability or making inferences about a population construct based on data or observations collected from a representative sample (Cook and Campbell, 1979; Gall et al., 2003; Isaac and Michael, 1997; Muijs, 2004; Sproull, 2002). Construct and external validity differ in that external validity is concerned with extrapolating statistical results from a sample to a specific target population (Cook and Campbell, 1979; Shadish et al., 2002). Isaac and Michael (1997) have identified 10 construct validity threats. To minimize the threat to construct validity arising from inadequate explication of constructs, this study was derive from a single straightforward construct (Cook and Campbell, 1979; Isaac and Michael, 1997). The construct tested was the instructional impact of an educational treatment on a
sample of university undergraduate students. The educational treatment was an instructional simulation with which all study participants interacted. The instructional impact was demonstrated, or not demonstrated, by performance on content-specific tests, and by knowledge seeking activities. Instructional impact will be measured by a statistical analysis of test scores. Construct validity threats arising from mono-operation bias were addressed by looking at two exemplars, course grade and knowledge seeking behavior (Cook and Campbell, 1979; Isaac and Michael, 1997). Course grade was based on a summation of all pretest and posttest grades as well as lab and quiz grades for each participant. Knowledge seeking behavior was assessed based on data collected from a survey instrument created for this study. Mono-method bias, utilizing a single measure of the dependent variable, was minimized by running a series of statistical tests on the data including analysis of variance (ANOVA) and multiple linear regression (Cook and Campbell, 1979; Isaac and Michael, 1997). Threats arising from hypothesis guessing could not be completely controlled due to the necessity of providing study participants with informed consent (Cook and Campbell, 1979; Isaac and Michael, 1997). However, at no time during the study did I meet or interact with study participants. In addition, the instructor of GNSC 1204 did not provide study participants with information about the study other than what was necessary to comply with MSU’s HSRC requirements. As study participants were predominantly Education majors it was perhaps inevitable that some of them should guess the study hypothesis. When study participants guess or learn to much about the purpose of the research, some may demonstrate evaluation apprehension in which participants engage in extraordinary activities to positively or negatively influence study results (Cook and
Campbell, 1979; Isaac and Michael, 1997). While providing participants with limited information about the study helped to reduce the evaluation apprehension threat, the necessity of informed consent precluded mitigation of the threat entirely. Threats arising from experimenter expectancies are perhaps one of the more challenging construct validity threats to address (Isaac and Michael, 1997). Two measures were utilized to circumvent this threat. First, the instructor who administered the assessments and collected the data was not involved in subsequent analysis of the data. Second, while I performed all data analysis, administering assessment instruments and data collection were carried out solely by the Associate Professor. Measures of this type have been documented in the literature as effective against experimenter expectancies (Cook and Campbell, 1979; Gall et al., 2003; Muijs, 2004). Construct validity threats arising from confounding constructs or levels of constructs arise when a continuous independent variable is measured on more than one level or when two continuous variables are not linearly related (Cook and Campbell, 1979; Isaac and Michael, 1997). To minimize this threat, the criterion variable in this study was measured at only one level. In addition, all scores measured were continuous variables which were linearly related along their continuums. The interaction of different treatments threat did not arise because participants in this study experienced a single treatment (Cook and Campbell, 1979; Isaac and Michael, 1997). Interaction of testing and treatment or pretest / posttest sensitization can arise when a pretest sensitizes study participants to the treatment prior to the treatment being administered or when a posttest provides hints or clues to interpreting the treatment (Cook and Campbell, 1979; Gall et al., 2003; Isaac and Michael, 1997). This interaction can impact posttest scores as well as the overall study
outcome and is also a threat to ecological validity. The Associate Professor’s selection of test questions was focused towards mitigating the interaction of testing and treatment threat.

Ecological validity has been defined as the degree to which study results can be extrapolated from an experimental setting to a natural or real-world setting (Gall et al., 2003). The most reliable method for circumventing this threat as documented in the literature is to administer the treatment and posttest without pretest then compare these results to another group which received the pretest. The study venue did not permit use of subgroups in order to address this issue.

The final threat to construct validity, that of restricted generalizability, is perhaps the most difficult to comprehensively address. It was impractical and beyond the scope of this study to address every venue to which the study construct might be extrapolated. In planning for the study, consideration was given to sample selection, study venue and data analysis methods with regard to generalizability of study results. However, the only reliable method for evaluating generalizability and ecological validity is replication of the study in other non-experimental settings (Cook and Campbell, 1979; Gall et al., 2003).

External Validity

External validity has been defined as the representativeness or generalizability of study results to a similar and/or larger population or different venue (Campbell and Stanley, 1963; Cook and Campbell, 1979; Gall et al., 2003; Isaac and Michael, 1997; Sproull, 2002). External validity addresses the application of the study treatment (X) to enquiries beyond the study scope as well as those variables, measurements and findings that can be extrapolated to other settings and populations. Three threats to
external validity have been identified by most authors (Campbell and Stanley, 1963; Cook and Campbell, 1979; Gall et al., 2003; Isaac and Michael, 1997). External validity threats from interaction of selection and treatment can arise when the study treatment cannot be generalized beyond the sample under study (Campbell and Stanley, 1963; Cook and Campbell, 1979; Gall et al., 2003; Isaac and Michael, 1997). The study population, undergraduates in public universities, the venue, Midwestern State University, and the sample, MSU undergraduate students enrolled in a general science course for non-science majors, were selected to facilitate the representativeness or generalizability of study results. The venue and sample selected should provide good population and ecological validity (Gall et al., 2003). See Appendix G for a detailed demographic of the sample and population selected for this study. Likewise the treatment, an Internet-based science simulation, was selected because of its very high degree of availability to students in a wide range of settings. These same measures also served as proofs against threats arising from interaction of setting and treatment as well as interaction of history and treatment.

Hawthorne Effect

The Hawthorne effect has been defined as a phenomenon which arises in research studies when study participants briefly alter their performance or behavior. This change in behavior is usually in response to increased attention from the researcher. As a consequence of MSU’s HSRC requirements for informed consent on the part of all participants in this study, the Hawthorne effect could not be completely circumvented (Cook and Campbell, 1979; Gall et al., 2003; Isaac and Michael, 1997). However, efforts to minimize the effect were taken in two areas. First, with the exception
of the knowledge-seeking behavior questionnaire to be administered at the end of the semester, pretests, posttests and all other in-class activities resembled as closely as possible what students would have expected to encounter in the absence of this study.

Second, while my name did appear as author of this study on the HSRC informed consent document presented to all participants, at no time did I interact with study participants. See Appendix C.

Selection of Samples

The population for this study consisted of 3,568 full-time undergraduate students enrolled at MSU during the spring 2007 term. The student population demographic of MSU for the spring 2007 term was composed of 58% women, 42% men, 64% White non-Hispanic, 15% African American, 4% Asian American or Pacific Islander, 8% Hispanic American, 1% Native American, 7% international representing 34 other countries and 1% not reported by student. See also Appendix G. No subpopulations were utilized for either the pilot study or the full study. Both the pilot study \((n = 68)\) and full study \((n = 78)\) utilized a convenience sample of participants from the MSU population. These participants were predominantly first and second-year undergraduate students from the West College of Education. The pilot study sample was composed of students enrolled in GNSC 1104 Life/Earth Science during the fall 2006 term. GNSC 1104 is offered by the Department of Science and Mathematics and is part of the core curriculum for all Education majors. For the full study the sample consisted of students enrolled in GNSC 1204 Physical Science. Once again, this course is offered by the Department of Science and Mathematics and is part of the core curriculum that all Education majors are required to complete. The large sample size for both the pilot
study \( n = 68 \) and the full study \( n = 78 \) strengthened reliability and internal validity of the study, reduced the probability of Type II errors arising from sampling and increased the power of statistical tests used (Isaac and Michael, 1997; Sproull, 2002).

Reliability has been defined as the degree to which test scores are free from error, as consistency of measurement or the degree to which study results can be consistently replicated (Muijs, 2004; Sproull, 2002). Internal validity has been defined by Gall et al. (2003) as the degree to which the effect on the research outcome of variables other than the treatment variable can be controlled or minimized. In drawing a sample from a population, the sample will deviate from the population to some degree due to the presence of chance variables (Isaac and Michael, 1997). Sampling error can diminish the generalizability of study results back to the parent population. Sampling error for this study was estimated using the standard formula (Hinkle et al., 2003; Sproull, 2002). With \( n = 78 \), the standard formula yielded a sampling error estimate \( e \) for the full study of ±0.111 given a confidence interval \( Z = 1.96 \) (95\%), \( p \leq .05 \). An \( \alpha \) value of .05 for both the pilot study and the full study contributed to statistical conclusion validity by minimizing threats arising from low statistical power. Power is expressed as a percentage and has been defined as the probability of correctly rejecting the null hypothesis when it is false thereby avoiding a Type II error (Cook and Campbell, 1979; Montgomery, 1984; Sproull, 2002).

**Procedures**

A pilot study was conducted during the fall 2006 semester, August 28 through December 16. The principle study was conducted during the spring 2007 semester.
The spring 2007 term at MSU began Monday, January 13. As part of the first or second class meeting the Associate Professor distributed informed consent forms in accordance with UNT IRB and MSU HSRC requirements. During the first weeks of the semester students completed a unit focusing on motion and force including a pretest-posttest. This unit constituted the first of three controls for this study. Subsequently students completed a unit focusing on sound wave propagation including a pretest-posttest. Curriculum for this unit centered on an instructional simulation for physical science, Properties of Mechanical Waves. A pretest was administered after which students were asked to complete a simulation exercise. This constituted the treatment phase of the study. A third pretest-posttest was administered focusing on electricity. A fourth pretest-posttest focusing on the periodic table was subsequently administered. These two pretest-posttests constituted the remainder of the control for this study. During the last week of the semester, on or before May 4, 2007, all participants were asked to complete a brief self-reporting questionnaire to collect data on knowledge-seeking behaviors. The survey instrument consisted of 11 questions and was created using the SurveyMonkey® online survey creation software (Finley, R., SurveyMonkey.com, Portland, OR, www.surveymonkey.com).

Pretest and posttest scores, daily quiz scores, knowledge-seeking survey results, attendance and student demographic data for the pilot study were collected and provided to me after December 15, 2006. Similar data from the full study were collected and delivered to me after May 14, 2007. Data sets I received were filtered by the MSU Assistant Director of Institutional Research to ensure that participant names, identification numbers or other personally identifying information were removed from the
Data analysis was carried out using the Statistical Package for the Social Sciences® (SPSS) (SPSS Inc., Chicago, IL, www.spss.com).

**Data Analysis**

Data analysis began shortly after the last day of the Spring 2007 term, May 14 and incorporated a number of statistical procedures. The same analysis procedures were carried out on data from the pilot study as well as data from the full study. See Appendix H.

**Statistical Methods**

Pearson product moment correlation was computed to ascertain the presence or absence of correlations between 10 predictor variables and course grade. Pearson r values were examined to test the hypotheses of significant correlation between treatment posttest scores and course grade.

A linear regression was computed to evaluate the amount of variance each of 10 predictor variables contributed to the criterion variable, course grade. This test also indicated which variables were the best predictors of the criterion variable. The linear regression tested the hypothesized relationships between the predictor variables and the criterion variable.

**Protection of Human Subjects, Sponsorship, and IRB Approval**

Participants in this study were all undergraduate students at MSU, Wichita Falls, Texas. All participants were at least 18 years of age. All participants were enrolled in GNSC 1104 Life/Earth Sciences for the pilot study and GNSC 1204 Physical Science
for the full study. Both courses are for undergraduate non-science majors and are offered by the Department of Science and Mathematics at MSU. GNSC 1104 and GNSC 1204 are part of the core curriculum requirements for all undergraduates in the West College of Education at MSU. Participants were informed of the purpose of the study by the Associate Professor. An informed consent statement explaining voluntary participation and the right to withdraw from the study at any time was distributed to all participants. The Associate Professor clearly stated that no psychological or physical harm to any participants were expected as a result of their involvement in the study.

To protect the anonymity of participants student names, social security numbers, university identification numbers and any other identifying personal data were removed from the data prior to analysis. I received no data until after the close of the fall 2006 semester. May research was sponsored by an Associate Professor of Education and Reading, West College of Education, MSU.

The IRB at the University of North Texas approved the research proposal for the present study. See Appendix B. The HSRC at Midwestern State University has also granted approval for this study. See Appendix D and F. I have completed the Human Participant Protection Education for Research Teams course from the National Cancer Institute as mandated by the IRB of the University of North Texas. See Appendix A.

Data collection began after August 26, 2006 and continued through May 10, 2007. Data generated by this study may be used for further research and publication.
CHAPTER 4

FINDINGS

Demographics

The population for this study consisted of 3,568 full-time undergraduate students enrolled at Midwestern State University (MSU) during the spring 2007 term. The student population demographic of MSU for the spring 2007 term was composed of 58% women, 42% men, 64% White non-Hispanic, 15% African American, 4% Asian American or Pacific Islander, 8% Hispanic American, 1% Native American, 7% international representing 34 other countries and 1% not reported by student. See also Appendix G.

Statistical Tests

Data

Of 78 observations, 5 observations (6.4%) were missing more than one data point. However, as these 5 observations retained more than half of the data points recorded, no observations were discarded from the data.

The gender variable was recoded in the Statistical Package for the Social Sciences® (SPSS) (SPSS Inc., Chicago, IL, www.spss.com) from string to numeric with the values F to 0 for female and M to 1 for male changed to 0 and 1 respectively. Of 78 participants, 6 were male and 72 were female. The class variable representing MSU student classification was recoded in SPSS from string to numeric.

When an ACT® college entrance and placement exam (ACT Inc., Iowa City, IA www.act.org) score / SAT® Reasoning Test (Educational Testing Services, Princeton, NJ, www.ets.org) score was present without an SAT score / ACT score, the
corresponding SAT score / ACT score was inserted into the dataset to replace the missing value. Replacement scores were taken from a standardized ACT-SAT conversion table, Appendix F. A total of 17 missing SAT scores (22%) were replaced by their corresponding ACT values by using the conversion table. See Appendix F. A total of 16 missing ACT scores (21%) were replaced by their corresponding SAT values by using the conversion table.

When both ACT and SAT scores were not present, the mean ACT score was used to replace the missing ACT score and the mean SAT score was used to replace the missing SAT score (Leech et al., 2005; Montgomery, 1984). The missing ACT score was obtained from a standard conversion table, Appendix F. This replacement occurred 21 times representing 27% of observations.

For the variable hsgpa (high school cumulative grade point average), 41 of 78 (52.6%) values were missing from the dataset. Rather than attempting to use a replacement method to estimate missing values resulting in potentially serious validly issues, this variable was not be used in the current analysis.

The variables for pretests, posttests, lab exercise, composite quiz grade, course grade, ACT score and SAT score where all measured on different scales. To obtain meaningful results it was necessary to perform a simple transformation to place all scores on the same scale. The transformation was accomplished by converting all scores to standard (z) scores using the following formula (Hinkel et al., 2003).

Pearson Product Moment Correlation

To ensure that all test scores were measured on the same scale, all scores were converted to standard z-scores prior to any statistical analysis. Test scores which
exhibited the most significant correlation with course grade were treatment posttest ($r = .763$), laboratory exercise grade ($r = .712$) and daily quiz grade ($r = .810$). A moderately significant correlation was found between the second control posttest and course grade ($r = .571$). All these correlations were significant with $p$ values less than .001. Given the high correlation between the treatment posttest and the variables for daily quiz, lab exercises and course grade, it is clear that the treatment had a positive impact on performance. High scores on the treatment posttest were significantly correlated with high lab exercise scores, high daily quiz scores and high final course grades. These findings supported the research hypothesis in that they indicate increased involvement with course content as reflected by improved grades.

*Linear Regression*

A linear regression was computed to evaluate the degree to which each of the 10 predictor variable contributed to the criterion variable, course grade. The 10 predictor variables included 4 sets of pretest-posttest scores, a laboratory exercise grade and a daily quiz grade. The first, third and fourth pretest-posttest sets served as controls while the second pretest-posttest covered the treatment phase of the study. Among the 10 predictor variables, 5 were significantly correlated with course grade; control pretest A1, treatment posttest B2, control posttest C2, lab grade and daily quiz grade. The other 5 predictors were moderately correlated with course grade.

Residuals from the analysis of variance (ANOVA) model summary table indicated a multiple $R$ of .983, an $R^2$ (effect size) of .966 and an adjusted $R^2$ of .960. The ANOVA model summary data indicated a strong fit with adjusted $R^2 = .960 > .05$. Minimal shrinkage (.006) between $R^2$ and adjusted $R^2$ indicated very little sampling error. These
results support the hypothesis that the sample was representative of the population with 96.6% of the variance in course grade explained or predicted by variance in the 10 predictors. A percentage of effect size was the result of error (e), also referred to as the coefficient of alienation (1 – 0.966 = 0.034). The error was small lending further support to the predictive validity of the predictor variables.

Results from the ANOVA summary table indicated an $F$-calculated of 187.967 with df(10, 67) and $\alpha = .05$. An $F$-critical value of 1.99 was obtained from a standard $F$-distribution table, Given that $F$-calculated > $F$-critical (187.967 > 1.99) once again the null hypothesis was rejected. A relatively large $F$-calculated value indicated that results were increasingly atypical.

From the ANOVA summary table an $\eta^2$ (correlation ratio) value was computed which indicated that 96.56% of the variance in the criterion variable course grade was explained by the variance in the 10 predictor variables. This large value for $\eta^2$ indicated a strong effect and significant statistical power (Hinkle et al., 2003; Leech et al., 2005; Muijs, 2004).

Significance values from the coefficients table indicated the statistical significance of relationship between each predictor and the criterion variable (Leech et al., 2005). At $p < .001$ only the treatment posttest, daily quiz grade and lab exercise grade had statistically significant relationships with course grade.

An examination of $B$ coefficients and $\beta$ weights clearly indicated that the treatment posttest, quiz grade and lab exercise grade contributed most significantly to course grade. At $\alpha = .05$ significance level, results for treatment posttest, quiz grade and lab exercise grad were significant at less than .05.
A significant regression equation was found \((F(10,67) = 187.967, p < .001)\), with an \(R^2\) of .966. The regression procedure resulted in the following regression equation.

\[
Y' = \alpha + \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n
\]

\[
Y' = (19.203) + (0.045*a1) + (0.011*a2) + (0.005*b1) + (0.399*b2) + (0.008*c1) + (0.027*c2) + (-0.006*d1) + (0.023*d2) + (0.238*quiz) + (0.344*lab)
\]

\(Y'\) (Y-hat) is referred to as a latent, synthetic or unobserved variable because it is created from a formula, not based on observation (Muijs, 2004). A synthetic variable \((Y^n)\) was computed for each observation in the data. The sum of \(Y^\wedge\) and residual random error variable \((e)\) for each observation was equivalent to \(Y\) (course grade) which validated the replicability of the results.

Linear regression results supported the research hypothesis. Students who performed well on the treatment posttest also performed well on lab exercises, daily quizzes and in the course overall. The treatment had a clearly measurable impact on student performance.

*Survey of Knowledge-seeking Behaviors*

A survey was administered to students enrolled in GNSC 1204 near the end of the spring 2007 term. The survey collected data on student’s knowledge-seeking behavior before and after interacting with the treatment. The treatment was the Properties of Mechanical Waves® simulation (Knight, San Francisco, CA : Addison-Wesley, wps.aw.com/aw_knight_physics_1/0,8722,1123696,-,00.html) a physical science simulation which presented concepts from sound wave propagation. The survey instrument was created using the SurveyMonkey® online survey creation software (Finley, R., SurveyMonkey.com, Portland, OR, www.surveymonkey.com). See Appendix.
E. A similar survey was administered to students in GNSC 1104 at the conclusion of the fall 2006 term as part of a pilot study. See Appendix H for a detailed description of the pilot study. The pilot study survey was administered during the final week of the semester and received a low response rate (27.94%). This low response rate rendered data collected from the survey anecdotal at best and of minimal analytical value. During the final week of the fall 2006 term students were concentrating on final exams and the completion of semester projects. Many students may have viewed the survey as just one more activity consuming valuable study time. The survey administered during the spring 2007 term was made available to students three weeks before the end of the semester. The result was a significant response (60.3%).

Student interaction with the Properties of Mechanical Waves simulation constituted the treatment phase of the study. Survey data were analyzed to address two hypotheses. First, did interaction with the simulation motivate students to undertake knowledge-seeking behaviors outside of class and beyond the course curriculum? This question addresses whether or not interaction with the simulation positively impacted student motivation to interact with course content. Second, did students acquire skills from the simulation which they were able to utilize during the subsequent control phases of the study? This question addresses whether or not students experienced skills and knowledge transfer as a result of interacting with the simulation.

The survey of knowledge-seeking behavior asked students to indicate if they had participated in one or more of 11 different knowledge-seeking behaviors during the one treatment and three control phases of the study. For a list of knowledge-seeking behaviors from the survey see Appendix E. During the three control phases of the study...
course content focused respectively on motion and force, electricity and, the periodic table. Course content during the treatment phase focused on sound and waves. As the treatment students interacted with the Properties of Mechanical Waves® simulation. Survey data indicated that 65.96% of respondents participated in at least one form of knowledge-seeking behavior. The percentage of students who participated in more than one knowledge-seeking behavior is also worthy of note: searched the internet, 59.6%; read all or part of a book [other than the course text], 42.6%; watched a science documentary, 29.8%; spoke with their professor outside of class, 29.8%. During the treatment phase of the study, 53.2% of students reported that they searched the Internet for additional information on sound and waves. This was an increase from 40.4% during the first control phase of the study. During the second and third control phases this percentage declined to 36.2% and 34.0% respectively. Also during the treatment phase, 34.0% of students reported having read all or part of a text, other than the assigned course text, on sound and wave propagation. This percentage represented a decline from 38.3% in the first control phase with percentages for the second and third control phases of 29.8% and 23.4% respectively. Student participation in other knowledge-seeking behaviors remained relatively constant or declined over the course of the semester.

Students were asked to indicate whether GNSC 1204 was their first experience with a science course at the post-secondary level. Responses indicated that 58.7% of students had previously taken at least one science course while 41.3% indicated that GNSC 1204 was their first experience in a post-secondary science course. It is reasonable to assume that a majority of students who reported taking science courses
prior to GNSC 1204 had previously taken GNSC 1104. Both courses are part of the core curriculum in the West College of Education at MSU, are taught by an Associate Professor of Education and Reading and the same text is used for both courses.

Students were asked to indicate the location where they worked with the simulation. The largest group, 59.6%, reported that they had worked with the simulation primarily at home or in their residence hall rooms. The remainder of respondents reported working with the simulation in a university computer lab (34.0%) or in some other location (10.6%).

With regard to the question of whether interaction with the simulation motivated students to undertake knowledge-seeking behaviors outside of class, the response rates of 53.2% for searching the Internet and 34.0% for reading a book other than the course text lends support to the hypothesis that simulations motivate students to interact with instructional content. Although these percentages declined during the second and third control phases of the study, they remained significant throughout the semester yielding semester total percentages of 59.6% (Internet searching) and 42.6% (reading a book other than the text). The percentage of students who reported speaking with their professor about course content outside of class declined during the second control phase and rose during the third control phase finishing up with a semester percentage of 29.8%.

With regard to the question of whether students acquired skills from the simulation which they utilized during the subsequent control phases of the study, no definitive conclusion can be drawn from the survey data.

Student responses to open-ended survey questions yielded interesting subjective
results. When students were asked to describe in their own words what they learned from the simulation, 76.6% expressed having gained new insights or better understanding of concepts from sound and wave propagation. Students also commented that the simulation aided in their understanding by allowing them to immediately see the impact on cause and effect relationships of manipulating variables. The same percentage of students reported that the simulation enabled them to learn differently by allowing them to visualize, understand and grasp concepts from wave propagation. Students commented positively about the hands-on experience the simulation provided and that the simulation facilitated learning how something works rather than just learning about a concept.

Students were asked to describe, in their own words, what they liked, or did not like, about interacting with the simulation. Most respondents, 89.4%, reported that they liked the experience of interacting with the simulation. In particular, students liked the interactivity and engagement the simulation provided. Only 10.6% of students reported that they did not like using the simulation. These students commented that some parts of the simulation graphics were difficult to understand while others reported technical problems running the simulation on their computers.

When asked to express, in their own words, the effect using the simulation had on how they studied electricity (control phase two) and the periodic table (control phase three) 61.7% of students reported that the simulation increased their general understanding of science and helped them approach the study of scientific concepts. Most students reported that the simulation was easy to use, that they enjoyed manipulating simulation variables and observing the results. In addition, 11.1% of
students reported searching the Internet for similar simulations to aid in their studies.

Summary

Throughout this study reasonable steps have been taken to strengthen aspects of the study methodology susceptible to threats arising from statistical conclusion validity, internal validity, face validity, construct validity, ecological validity and external validity. Results of statistical tests were evaluated from a conservative perspective in the interest of limiting the probability of Type I and Type II errors. A target significance level of $\alpha = 0.5$ was selected prior to data collection and statistical testing.

The hypotheses tested in the study was whether transfer of knowledge and skills from an instructional setting in which simulations are used to a similar instructional setting where simulations are not used can be demonstrated to have occurred. An examination of measures of central tendency as well as a linear regression computation yielded significant results. While some distributions were moderately platykurtic or exhibited a small degree of skewness, all distributions were sufficient normal to preclude any transformations.

Pearson coefficients and linear regression results indicated significant correlation between the treatment posttest predictor variable, quiz grade and lab exercise grade, and the criterion variable, course grade. These results supported the research hypothesis. Participant responses to a knowledge-seeking behavior survey indicated that 65.96% of respondents participated in at least one of 11 knowledge-seeking behaviors. Survey responses were predominantly positive (89.4%), lending further support to the research hypothesis.
The first successful classroom use of an instructional simulation took place in the Harvard Business School in 1909 (Saunders, 1997). Classroom use of simulations is expanding but remains limited. This limitation is to a large degree a consequence of problematic research studies (Washbush and Gosen, 2001). Much extant research suffers from poor research designs, limited sample size, inappropriate data analysis techniques or lack of informative data reporting (Akpan, 2001). The goal of this study was to mitigate the limitations of previous research while investigating the learning effectiveness of integrating a simulation into a traditional university course curriculum. The study venue was one section of an undergraduate general-science course for non-sciences majors which provided a convenience sample of $n = 78$.

Much extant research indicates that simulations possess great potential as aids to instruction (Aldrich, 2005; Forinash and Wisman, September, 2001). The purpose of this study was to further document the significant role simulations can play in the classroom as providers of opportunities for situated learning experiences. The questions addressed by this study focused on whether simulations contribute to instruction by facilitating transfer, improved motivation and increased engagement. Further studies of this type are warranted in order to gain a more complete understanding of how students are impacted by their interactions with simulations as well as the role simulations can play in the classroom.

The hypothesis tested was whether transfer of knowledge and skills from an instructional setting in which simulations are used to a similar instructional setting where
simulations are not used can be demonstrated to have occurred. Descriptive and frequency statistics were analyzed to ascertain whether the variables under study represented normal distributions. All distributions were sufficiently normal to the extent that transformations were performed other than converting all values to z-scores.

Pearson r correlations and linear regression were computed to look for significant correlations as well as determine how well the criterion variable, crsgrad, was predicted by 10 predictor variables. Analysis of variance (ANOVA) results indicated minimal shrinkage, minimal sampling error, significant effect and good statistical power. Based on β coefficients the treatment posttest, daily quiz and lab grades had statistically significant relationships with the criterion variable, course grade. Students who performed well on the treatment posttest also performed well on lab exercises, daily quizzes and in the course overall. Clearly the presence of the instructional simulation in the curriculum had a significant positive effect on the learning experience.

A survey was administered to study participants which assessed knowledge-seeking behaviors. Participant response was 60.3%. Survey data indicated that 66% of respondents participated in at least one of 11 different forms of knowledge-seeking behavior. Survey responses were predominantly positive. These results lend further support to the research hypotheses and indicate that the treatment, an instructional simulation did facilitate a positive learning experience.

Theoretical Consequences

This study presented theoretical consequences in regard to motivation and engagement. Sufficient evidence emerged from this study to support claims that simulations do in fact promote discovery-learning and motivate students to interact more
deeply with instructional content. Further, the positive impact of interacting with the simulation does exhibit longevity and the motivation, if not skills, instilled by interacting with the simulation does transfer. The validity of Vygotsky’s zone of proximal development (ZPD) is also demonstrated by this study in that students who experienced the simulation were motivated to stretch and seek more information above and beyond the boundaries of course content.

Methodological Issues

This study has demonstrated the methodological weaknesses inherent in most current efforts to examine the impact of simulations on learning. While evidence of how simulations impact learning can be observed through analysis of changes in scores, this approach still leaves many variables unaddressed. Specifically, which aspects of simulations have the most positive and enduring impact on students remains an open question. Further, there is the question of whether students really do view simulations as video games or as something significantly different. While there is evidence to support claims that simulations provide situated learning opportunities, no systematic approach exists to quantitatively and exhaustively explore and test this assumption. Finally, while the number of instructional simulations for instructors to choose from is growing continuously, resources to assist instructors with integrating simulations in their classroom activities remain woefully limited. Efforts towards development of systematic assessments of simulation effectiveness will be significantly restricted so long as instructors view simulations as specialized tools not for general classroom use. At the same time, until simulations achieve wider use data collection and analysis will remain problematic. In the absence of more meaningful studies built on sound theoretical
foundations, instructors have few sources to aid in evaluating the possible benefits of simulations for their students.

The Need for Further Research

Research and investigation of design and assessment of instructional simulations is crucial because simulations hold substantive promise for learning. The results of this study have indicated that simulations can motivate students to engage as well as take a more active role in their own learning. While innovative and intriguing research is currently ongoing, future research efforts should be focused on several specific areas.

- Quantitative data on the efficacy of classroom use of simulations is limited. Much of the extant research exhibits weaknesses in measurement and design. All too often studies do not describe the post-tests employed or how these tests effectively measure learning outcomes. Student characteristics, such as motivation, which may interact with instruction, are often overlooked when collecting data. This omission ignores the fact that not all students will react in similar or consistent ways to simulations. Additional studies need to be conducted using larger samples, sound experimental designs and reliable measurement procedures.

- Far too many research projects have been undertaken which merely compare the use of simulations to conventional teaching methods. The lack of meaningful data such studies can reveal has been well documented in the literature.

- Data on the use of simulations specifically designed to facilitate situated learning and anchored instruction is limited further still. Research projects
focusing on the efficacy of simulations incorporating multi-sensory 3D environments are also severely lacking.

- The roles of feedback, coaching, scaffolding, learner control and debriefing in simulations as well as the roles of motivation and competition warrant further investigation.

- Development of a comprehensive simulation design paradigm based on sound learning theory is lacking. These research efforts should focus on the development of a comprehensive authoring environment which can be used by instructional designers, content experts and practitioners with the necessity of confronting a steep learning curve.

- Researchers often fail to document observations of student interactions with their classmates and with the content materials while using simulations.

- The design and development of artificial intelligences and agents to play the roles of mentors, allies or adversaries in simulations is yet another area in serious need of research (Aldrich, 2005; Evans, 2001).

- Further quantitative study of the various simulation models and their support for specific learning styles and strategies, as well as development of new models, is necessary (Saunders, 1997).

- Little investigation has been made into the design and application of simulations to promote generative learning in which students confront contradictory ideas and resolve their own inner conflicts.

- Further investigation is needed into how design of simulations can provide intuitive, widely applicable anchored learning experiences.
As bandwidth becomes less of a constraint, Internet access becomes ever more ubiquitous and the price of technology continues to fall, educators can expand their utilization of instructional technologies like simulations both in and outside the classroom. However, without a clear understanding of how best to use simulations, we have no assurances that they can make a positive contribution to instruction. To make effective decisions on how and where to best integrate simulations into their curriculum, educators need data from substantive, quantitative studies. A central goal of this study was to provide additional quantitative data upon which informed decisions can be based. This goal was achieved.

Because simulations provide environments in which the learner can construct new knowledge, they stand poised to make a significant contribution to education. The benefits of “high-quality learning” which simulations can provide to the student are well documented in the literature (Forinash and Wisman, September, 2001; Aldrich, 2005). However, substantive research studies, similar to that documented in this dissertation, are essential in assessing any possible instructional benefits of simulations as modalities for situated learning.

Conclusions

The purpose of this study was to test the hypotheses that simulations can motivate students to engage with instructional content and that students will subsequently experience transfer of knowledge and skills. Through the use of quantitative and qualitative methods this study has documented that simulations can have a positive impact on learners and that this impact can be observed in test scores. This positive impact was also manifested in student motivation to interact with course
content. After interacting with the simulation students demonstrated engagement in extra-curricular activities in an effort to learn more on their own. In addition, students used the Internet to search for simulations indicating that the simulation experience was both positive and had lasting efficacy.

While this study has produced significant results, much more could be accomplished by replicating this study as a true experiment. This would require a venue in which two sections of the same course with like numbers of students in each section could be designated for the experiment. The same instructor would teach both sections utilizing the same text and curriculum in both classes. The section designated as the treatment would interact with an instructional simulation at some point during the term while the control section would not use a simulation. Given an adequate sample size and a sufficient number of scales, significant results could be anticipated from such an experiment.
APPENDIX A

NATIONAL CANCER INSTITUTE COURSE COMPLETION CERTIFICATE
In compliance with 45 CFR Part 46 Section 46.101, the Institutional Review Board (IRB) of the University of North Texas (UNT) requires that all researchers who will conduct studies involving human participants must complete a course in human participant protection. The course used by UNT's IRB is currently offered over the Internet by the National Cancer Institute and is titled, Human Participant Protection Education for Research Teams. Researchers who successfully complete the course receive a certificate a copy of which must be presented to the IRB at UNT. This appendix contains a reproduction of the completion certificate for the author of this study.
Completion Certificate

This is to certify that

Les Lunce

has completed the Human Participants Protection Education for Research Teams online course, sponsored by the National Institutes of Health (NIH), on 07/01/2006.

This course included the following:

- key historical events and current issues that impact guidelines and legislation on human participant protection in research.
- ethical principles and guidelines that should assist in resolving the ethical issues inherent in the conduct of research with human participants.
- the use of key ethical principles and federal regulations to protect human participants at various stages in the research process.
- a description of guidelines for the protection of special populations in research.
- a definition of informed consent and components necessary for a valid consent.
- a description of the role of the IRB in the research process.
- the roles, responsibilities, and interactions of federal agencies, institutions, and researchers in conducting research with human participants.

National Institutes of Health
http://www.nih.gov
APPENDIX B

UNIVERSITY OF NORTH TEXAS (UNT)

INSTITUTIONAL REVIEW BOARD (IRB) APPLICATION FOR REVIEW
In compliance with 45 CFR Part 46 Section 46.101, the Institutional Review Board (IRB) of the University of North Texas (UNT) requires that all researchers who will conduct studies involving human participants must submit an application for approval of research. Once the IRB is assured of compliance by the researcher, a letter of approval or exemption from further review is issued. This appendix contains a reproduction of the completed IRB application review form and the letter of exemption of further review.
April 18, 2007

Les Lunce
Department of Technology and Cognition
University of North Texas

RE: Human Subjects Application No. 07-143

Dear Mr. Lunce:

In accordance with 45 CFR Part 46 Section 46.101, your study titled “An Investigation of the Use of Instructional Simulations in the Classroom as a Methodology for Prompting Transfer, Engagement and Motivation” has been determined to qualify for an exemption from further review by the UNT Institutional Review Board (IRB).

No changes may be made to your study’s procedures or forms without prior written approval from the UNT IRB. Please contact Sheila Bourns, Research Compliance Administrator, ext. 3940, if you wish to make any such changes.

Sincerely,

Scott Simpkins, Ph.D.
Chair
Institutional Review Board

SS:3b
Application for Initial Review

University of North Texas Institutional Review Board
CHRP Federalwide Assurance: FW00007479

Please save this Application as a Word document on your computer, answer all questions completely, and submit it along with all supplemental documents to the UNT Office of Research Services as described on the Signature Page. Handwritten forms will not be accepted.

1. Principal Investigator Information
   Must be the same Principal Investigator named in any proposal for external or internal funding.

<table>
<thead>
<tr>
<th>First Name</th>
<th>Last Name</th>
<th>E-mail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee</td>
<td>Lunce</td>
<td><a href="mailto:lmunce@yahoo.com">lmunce@yahoo.com</a></td>
</tr>
</tbody>
</table>

If UNT faculty or staff:

<table>
<thead>
<tr>
<th>Department</th>
<th>Building</th>
<th>Room Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology &amp; Cognition</td>
<td>Matthews</td>
<td>313</td>
</tr>
<tr>
<td>UNT Department</td>
<td></td>
<td></td>
</tr>
<tr>
<td>940-565-2579</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Office Phone Number</td>
<td>Fax Number</td>
<td></td>
</tr>
</tbody>
</table>

If UNT student:

1409 Grant St.
Home address
Victoria Falls, TX 76109
City, State, and Zip Code
940-766-0485
Home Phone Number
3947-3122
Fax Number
jyoung@unt.edu
Dr. Jon Young

Comp. Ed. & Cognitive Systems, College of Education
Faculty Advisor’s Name
Faculty Advisor’s Office Phone Number
940-565-2057
Faculty Advisor’s UNT Department

Is this study for your master’s thesis or doctoral dissertation?
[ ] Yes  [ ] No

Is this study for your other course work?
[ ] Yes  [x] No

Under the UNT IRB Guidelines, research conducted solely for satisfying course requirements does not require IRB review for approval, unless the investigator intends to publish or publicly present the results of the study as "generalizable knowledge" for his/her field of study.

2. Co-Investigator Information
   If applicable, students should include their Faculty Advisor as Co-Investigator only if he/she is actively involved in conducting the study.

<table>
<thead>
<tr>
<th>First Name</th>
<th>Last Name</th>
<th>E-mail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bambi</td>
<td>Bailey, Ph.D.</td>
<td><a href="mailto:Bambi.Bailey@mwsu.edu">Bambi.Bailey@mwsu.edu</a></td>
</tr>
<tr>
<td>940-397-6237</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Office Phone Number</td>
<td></td>
<td>University or Other Entity</td>
</tr>
</tbody>
</table>

The Co-Investigator’s classification is:
[ ] Faculty/Staff  [x] Graduate Student

3. Key Personnel
   List the names of all other Key Personnel who are responsible for the design, conduct, or reporting of the study.

Form designed and maintained by UNT ORS, (940) 565-3940
Last Updated on 19 October 2005
4. Project Information

A project of the use of instructional simulations in the classroom as a methodology for promoting transfer, engagement and motivation.

Product Title (must be the same as any proposal for external or internal funding.)

An investigation of the use of instructional simulations in the classroom as a methodology for promoting transfer, engagement and motivation.

Start Date (mm/dd/year) 11/1/2006

End Date (mm/dd/year) 12/15/2006

Project Sponsors (identify the source(s) of any external and/or internal funding and attach a complete copy of the funding proposal.)

5. Significant Financial Conflict of Interest

If any external funding is proposed, have you and all Key Personnel submitted a Significant Financial Interest Disclosure form in compliance with the UNT Conflict of Interest Policy for Sponsored Projects? (For more information, see UNT Policy Number 16.12.3.3 at http://www.untpolicy/UNT_Policy/volume3/16.12.3.3.html.)

☐ Yes

☐ No

X No External Funding

6. NIH Training

Have you and all key personnel completed the required NIH training course (Human Participant Protection Education for Research Teams) and submitted a copy of the completion certificate to the Office of Research Services?

☐ Yes

☐ No

X No

If "No," this training is required for all key personnel before your study can be approved. This free on-line course may be accessed at http://cancancer.gov/clinicaltrials/learning/humanparticpant-protections.asp

7. Purpose of Study

In no more than half a page, briefly state the purpose of your study in lay language appropriate for the UNT IRB's community members and faculty members outside of your field; include the hypotheses or research question(s) you intend to answer; avoid cutting and pasting from any funding proposal, master's thesis, or doctoral dissertation; applications submitted with overly technical language will be returned to the Principal Investigator for revision before review by the IRB.
This study will analyze archival data collected by Dr. Bailey on behalf of Mr. Lunce during the Fall 2006 semester. The goal of this study is to investigate the impact instructional simulations have on the learning experience. The hypothesis to be tested are:

1. that instructional simulations increase student engagement with course content,
2. that instructional simulations provide a deep-learning experience and
3. that instructional simulations promote transfer of skills and knowledge to contexts outside the simulation.

This study will utilize a single-group pretest-posttest design with control. A content-based pre-test was administered to the sample at the beginning of the Fall 2006 semester. During the first half of the semester, students used an Earth-science simulation focusing on concepts from geology. Upon completion of the simulation, students completed a content-based post-test focusing on concepts introduced through the simulation. During the second half of the semester a second pre-test was administered focusing on topics from Ecology. Students used a standardized curriculum without simulation. A content-based post-test was administered at the end of the semester. Data was collected from exam scores and attendance records. A knowledge seeking survey was administered during the last week of the semester.

The population for this project were undergraduate students in the College of Education at Midwestern State University, Wichita Falls, Texas. A convenience sample of approximately 86 students from this population was selected. The sample consisted of primarily first and second-year students enrolled in an undergraduate science course for non-science majors, offered by the Department of Science & Mathematics. This was a core course which all Education majors must take.

8. Previous Research

Seminar on previous research leading to the formulation of this study, including any past or current research conducted by the Principal Investigator or key personnel.

The use of simulations for classroom instruction can be traced to the 1970's. However, research on the effectiveness of simulations in the classroom as aids to instruction remains limited and to some degree problematic. Much extant research suffers from poor research designs, inappropriate data analysis techniques or lack of informative data reporting. Accordingly, further experiments are warranted in order to gain a more complete understanding of how students are impacted by their interactions with simulations as well as the role simulations can play in the curriculum.

The goal of this study is to conduct such an experiment.

The questions addressed by this study focus on whether simulations contribute to instruction by facilitating transfer, improved motivation and increased engagement. This study will not compare instructional simulations to traditional classroom instruction as this would have constituted yet another media-comparison study. Such studies have been shown to produce little in the way of significant statistical results. The venue, sample size, experimental design and data analysis methods selected for this study were chosen specifically to contribute to the validity and reliability of obtained results.

9. Informed Consent Forms

Written Informed Consent Forms signed by the subject or the subject’s legally authorized representative are required for most IRB projects (exceptions include telephone surveys and internet surveys for which an Informed Consent notice is substituted). If any subjects will be children (under 18 years of age in Texas), the parent/guardian Informed Consent Form must include a section for obtaining assent by children ages 7-17. Submit a copy of all consent/assent forms to be used.

Templates for creating informed consent forms are located on the Office of Research Services website at the address shown for each type of study:

Faculty/staff investigators:
adult consent (ages 18 or older): http://www.unt.edu/ospa/docs/IRB_Consent_FSAS.doc
parent/guardian consent with minor assent (ages 7-17): http://www.unt.edu/ospa/docs/IRB_Consent_FSMA.doc

Student investigators:
adult consent (ages 18 or older): http://www.unt.edu/ospa/docs/IRB_Consent_GRAS.doc
parent/guardian consent with minor assent (ages 7-17): http://www.unt.edu/ospa/docs/IRB_Consent_GRMA.doc

IRB Application for Initial Review, Page 3
10. Foreign Languages

Will your study involve the use of Informed Consent Forms, data collection instruments, or recruitment materials in any language other than English?

[X] No

If "Yes," identify all foreign languages below. Please do not submit any foreign language forms or materials until after the IRB has approved the English versions.

11. Informed Consent

Describe the steps for obtaining the subjects’ informed consent/assent (by whom, where, when, etc.).

The purpose of the study was explained by Dr. Bailey to participants during the first or second class meeting, at the beginning of the Fall 2006 semester, on or after August 26, 2006. Participants were informed that an instructional simulation will be used during the first half of the semester as part of the regular course curriculum. Informed consent forms provided by Midwestern State University’s Human Subjects Research Committee (HSRC) were distributed to all participants. IFBG stipulated that introduction of consent forms or similar documents from other institutions was explicitly prohibited. Explanations of the study and participant questions regarding the study were addressed by Dr. Bailey.

12. Medications

Will any subjects be under the influence of any medication, drugs or stressful condition which could diminish their ability to give effective informed consent?

[X] No

If "Yes," please explain and describe what steps you will take to verify that potential subjects possess the mental capacity to give meaningful informed consent to participation in the study.

13. Location of Study

Identify all locations where the study will be conducted. For each data collection site other than UNT, attach a signed and dated original of a letter on the cooperating institution’s letterhead giving approval for collection of data at that site. This letter should reflect a general understanding of the nature of the study and how it will be conducted.

The study will take place during the Fall 2006 (8/26/06 — 12/15/06) semester on the campus of Midwestern State University (MSU), 3140 Taft Boulevard, Wichita Falls, Texas, 76308. A letter of approval has been received from MSU’s Human Subjects Research Committee. This letter will be provided to UNT’s IRB upon request.

14. Recruitment Population

Describe the population from which the subjects (including controls, if applicable) will be recruited.

The population for this project was restricted to undergraduate students in the College of Education at Midwestern State University, Wichita Falls, Texas. Research was conducted using a convenience sample of approximately 68 students from this population. The sample consisted primarily of first and second year students enrolled in an undergraduate science course for non-science majors offered by the Department of Science & Mathematics at MSU. This is a core course which all Education majors must take.

15. Subject Recruitment

Describe how you will recruit subjects to participate in the study; attach a copy of all recruitment materials (newspaper advertisements, posters, telephone scripts, etc.).

IRB Application for Initial Review, Page 4
To participate in the study, students had to be enrolled in GNSC 1104, an undergraduate science course offered by the Department of Science & Mathematics. No other recruitment procedures were undertaken.

16. Subject Population Composition
Describe the anticipated gender, race/ethnic composition, age range and health status of the study population and the criteria for inclusion or exclusion of any subpopulation.

The study population was 5,544 undergraduate students at Midwestern State University, a public university located in North-central Texas (Region 9). The student population demographic for 2006 was composed of 57% women, 43% men, 13% African American, 3% Asian American or Pacific Islander, 9% Hispanic American, 1% Native American, international population 5% representing 37 other countries. No subpopulations will be utilized.

17. Vulnerable Populations
Please identify any vulnerable populations who will specifically be targeted for participation in this study:

- [ ] Children (under 18 years of age)
- [ ] Pregnant women
- [ ] Prisoners, including juveniles
- [ ] Mentally impaired or mentally retarded

If any boxes are checked, describe any special precautions to be taken in your study due to the inclusion of these populations:

18. Number of Participants
Total number of subjects (including controls):

68

Number of controls (if applicable):

19. Data Collection
Describe all procedures you will use to collect data (interviews, surveys, focus groups, observation, review of existing records, etc.). Attach a copy of all data collection instruments and interview scripts to be used.

Data was collected form pre-test and post-test scores, and attendance records. A brief questionnaire on knowledge seeking behaviors was administered at the end of the semester.

20. Time
Estimate the total time each subject will be involved in the study (include time per session, total number of sessions, etc.).

The study will take place during the Fall 2006 (8/26/06 – 12/15/06) semester. Participants were enrolled in a 4-credit hour undergraduate science course for non-science majors. The amount of time each student was expected to participate in the study did not exceed the time a student would normally have spent preparing for and attending the class. No additional class meetings or interviews were conducted.

21. Compensation
Describe any payment or other compensation subjects will receive for participating in the study, including the timing for payment and any conditions for receipt of such compensation.

No payment or disbursement will be offered or permitted.
22. Risks and Precautions
Describe any foreseeable risks to subjects presented by the procedures described above in the Data Collection section, including any physical, psychological, social, economic, legal, or confidentiality risks (see the UNT IRB Guidelines for more information about these risks). Include your assessment of the degree of each risk presented and all precautions you will take to minimize such risks or to respond to any adverse events, should they occur.

No withdrawal from treatment or other risks are expected.

23. Benefits
Describe the benefits to the subjects or others (explain how the subjects will benefit from participating in the study, other than any compensation described in the Compensation section above; if the subjects will not directly benefit from the research, explain how the study will benefit others or contribute to your field of research):

Benefits to the human participants involved in the study.

Students will benefit from a better understanding, on the part of their instructors, of optimal utilization of instructional simulations in the classroom. A more focused and optimized integration of instructional simulations into the traditional curriculum will result in an enhanced learning experience for the student. In addition, meta-cognitive skills acquired by the student through interaction with a simulation can positively impact the student's learning experience in other domains.

Benefits to individuals who are not participants, but who may have similar problems/concerns:

Meta-cognitive skills acquired by the student as an outcome from interaction with an instructional simulation can positively impact the student's learning experience in other domains. Students will perform better in other university courses and more effectively transfer skills and knowledge acquired to real-world settings.

Benefits to society in general:

Meta-cognitive skills acquired through interaction with an instructional simulation can facilitate transfer of knowledge from a learning environment to real-world settings. Effective use of meta-cognitive skills in the workplace results in more productive employees. Employers will benefit from staff possessed of well-developed problem-solving skills. The ability of individuals to resolve ill-defined problems and effectively address open-ended scenarios will benefit society as a whole.
24. HIPAA

Will your study involve obtaining individually identifiable health information from health care plans, health care clearinghouses, or health care providers?  

☐ Yes  ☑ No

If "Yes," describe the procedures you will use to comply with the HIPAA Privacy Rule. (For more information about HIPAA, see the HIPAA Guidance page on the UNT Research Services website at http://www.unt.edu/research/hipaa.htm.)

25. Confidentiality of Research Records

Describe the procedures you will use to maintain the confidentiality of any personally identifiable data (including any videotapes and/or audiotapes of the participants).

No audio or video recordings of participants were made or utilized. Confidentiality of participant data during analysis will be maintained by assigning each participant a randomly generated identification number. Student names, student identification numbers, social security numbers, etc. will be removed from the data prior to analysis. Data will be viewed only by Dr. Bailey and Mr. Lunco.

Describe where your research records will be maintained, any coding or other steps you will take to separate participants' names from research data, and how long you will retain personally identifiable data in your research records.

Research data will be retained on computer used exclusively by Dr. Bailey and Mr. Lunco. These computers will be secured in each researcher's perspective offices. Password protection and encryption of data files will be utilized as needed. Participants will be assigned a randomly generated identification number to separate participants from personal data. Names, student identification numbers, social security numbers, etc. will not be retained in the data. Upon completion of data collection, all personally identifiable data will be discarded. Data collection is estimated to conclude on or shortly after December 15, 2006.

Identify the categories of all persons other than the research team to whom personally identifiable data will be disclosed and the purpose of each such disclosure (presentations at academic conferences, dissertation committee, etc.).

Participant personal data will not be disclosed to personnel other than members of the research team: Dr. Bailey and Mr. Lunco.

26. Publication of Results

Please identify all methods in which you plan to publicly disseminate the results of your study (academic journal, academic conference, thesis or dissertation, etc.).

Research results will be disseminated through defense and publication of dissertation by Mr. Lunco. In addition, research results will be presented for publication in professional journals authored jointly and/or singly by Dr. Bailey and Mr. Lunco. Results will be further disseminated through presentations and poster-sessions at professional conferences. Where appropriate, partial or summary results may be presented in electronic forums or journals on the Internet. Finally, research findings will be shared with colleagues on an informal basis when appropriate.
IRB Application Signature Page

Principal Investigator
I certify that the information in this application is complete and accurate. I agree to conduct this study in accordance with the UNT IRB Guidelines and the study procedures and forms approved by the UNT IRB. I agree that I will not make any changes to the approved procedures or forms without prior written approval from the UNT IRB. I understand that I cannot initiate any contact with human subjects until I have received written UNT IRB approval.

[Signature]

Signature of Principal Investigator
Date
July 5, 2006

Faculty Advisor (if applicable)
I have examined this completed application and I am satisfied with the adequacy of the proposed research design and the precautions to be taken for the protection of human subjects. My oversight of this study will include verification that it is being conducted in accordance with the UNT IRB Guidelines and the study procedures and forms approved by the UNT IRB. I agree that no changes will be made to the approved procedures or forms without prior written approval from the UNT IRB.

[Signature]

Signature of Faculty Advisor
Date

Submission of Your Application and Supplementary Documents
Print the entire application and sign this page. If you are a student, ask your Faculty Advisor to also sign this page. Attach all supplementary documents, including:
- Copies of all NIH Training completion certificates not previously submitted to the Office of Research Services;
- A copy of any proposal for internal or external funding for this study;
- The original of the approval letters from all cooperating institutions (other than UNT) where you will collect data;
- A copy of all recruitment materials;
- A copy of all informed consent forms; and
- A copy of all data collection instruments.
Send or deliver the entire application (including this Signature Page) and all supplementary documents to:
Physical Address:
Office of Research Services
Hurley Admin. Bldg. 160
Mailing address:
UNT Office of Research Services
P.O. Box 305250-5250
Denton, TX 76203
Thank you for submitting your application to the UNT IRB. Please contact Sheila Bourns at (940) 565-3940 or sbourns@unt.edu for any questions about your application.
In compliance with 45 CFR Part 46 Section 46.101, the Human Subjects Review Committee (HSRC) of Midwestern State University (MSU) requires that all researchers who will conduct studies involving human participants must submit an application for approval of research. Once the HSRC is assured of compliance by the researcher, a letter of approval or exemption from further review is issued. This appendix contains a reproduction of the completed HSRC application review form and the letter of exemption of further review.
MEMORANDUM

TO:  Les M. Lunce & Dr. Bambi Bailey

RE:  HSRC Application

DATE:  August 29, 2006

Please be advised that your application for research utilizing human subjects has been reviewed
and approved by the above named committee. The number assigned this project is:

File number:  06082901

Please include this number in any presentation or publication arising from this research. You
may be required to place a copy of this letter within the thesis or other class, department, or
college documentation. This approval is valid for one calendar year following granting of
approval status. Your may request an extension by submitting a letter requesting such to the
HSRC committee chair.

Respectfully,

[Signature]

Chair, Human Subjects in Research Committee
ATTACHMENT A
Midwestern State University
(For Office Use: Proposal Number 052911)

CLAIM FOR EXEMPTION FROM REVIEW BY THE HUMAN SUBJECTS REVIEW COMMITTEE

NOTICE: Collection of data may begin only after this form has received Committee approval and has been properly filed with the Chair, Human Subjects Review Committee. The Committee may, upon review of this claim, deny the request for an exemption.

To: Chair, Human Subjects Review Committee
Midwestern State University

1) Name(s) of Principal Investigator(s): Les M. Lunce & Dr. Bambi Bailey

Program Affiliation College of Education and Phone Number 940-765-0485, 940-397-6237

2) Title of Project or Proposal for which exemption is claimed:

An investigation of the use of instructional simulations in the classroom as a methodology for promoting transfer, engagement and motivation.

3) Description of Project or Proposal (attach additional information as required):

a. Briefly describe the population of human subject involved (e.g. University students, community members, athletes, homemakers, school children, etc.)

The population for this project are undergraduate students in the College of Education (COE) at Midwestern State University, Wichita Falls, Texas. Research will be conducted using a sample of approximately 100 students from this population. The sample will consist of primarily first and second year students enrolled in an undergraduate course. This course is one of the core science content courses within the College of Science and Mathematics that education majors take.

b. You MUST indicate if this participation is ☐ voluntary or ☐ not.

c. Briefly describe your research procedures and techniques of data collection (e.g., interview, questionnaire, test administration, observation of public behavior, etc.).

This study will utilize a time-series design. A content-based pre-test will be administered to the sample at the beginning of the Fall 2006 semester. During the first half of the semester, students will use an Earth-science simulation focusing on concepts from geology. Upon completion of the simulation, students will complete a content-based post-test focusing on concepts introduced through the simulation. A content-based post-test will be administered mid-semester. During the second half of the semester students will use a standardized curriculum without simulations. A content-based, comprehensive post-test will be administered at the end of the semester. Data will be collected from exam scores, attendance records and from a brief demographic questionnaire.

d. Briefly describe the objectives of your research (e.g., what hypotheses you are testing).

The goal of this study is to investigate the impact instructional simulations have on the learning experience. The hypothesis to be tested are:

1. that instructional simulations increase student engagement with course content,
2. that instructional simulations provide a deep-learning experience and
3. that instructional simulations promote transfer of skills and knowledge to contexts outside the simulation.

Attachment A
Claim for Exemption
Page 1 of 4
4) Describe Subject Recruitment:
   a. How will you recruit subjects? (You MUST submit verbatim copies of all letters, notices, advertisements, etc., with an outline of all oral presentations to be used.)

   □ Direct Person-to-Person Solicitation
   □ Telephone Solicitation
   □ Newspaper Solicitation
   □ Letters of Solicitation
   □ Other (explain)

   Students enrolled in a core science content courses within the College of Science and Mathematics will be selected to participate in the study. This course is a core requirement for all Education majors. Participation will be based on enrollment.

   b. List all criteria for including subjects:

   Students enrolled in a core science content courses within the College of Science and Mathematics taught by Dr. Bambi Bailey during the Fall 2006 semester.

   Students must be undergraduates at Midwestern State University.

   c. List all criteria for excluding subjects:

   Subjects not meeting the above two criteria will not be included in the sample.

5) Describe Subject Benefits and Costs:
   a. Indicate what, if any, benefits may accrue to each of the following: (Payment to research subjects for participation in studies considered a benefit.)

   (1). The human subjects involved.

   Students will benefit from a better understanding, on the part of their instructors, of optimal utilization of instructional simulations in the classroom. A more focused and optimized integration of instructional simulations into the traditional curriculum will result in an enhanced learning experience for the student. In addition, meta-cognitive skills acquired by the student through interaction with a simulation can positively impact the student's learning experience in other domains.

   (2). Individuals who are not subjects, but who may have similar problems/concerns:

   Meta-cognitive skills acquired by the student as an outcome from interaction with an instructional simulation can positively impact the student's learning experience in other domains. Students will perform better in other university courses and more effectively transfer skills and knowledge acquired to real-world settings.

   (3). Society in general:

   Meta-cognitive skills acquired through interaction with an instructional simulation can facilitate transfer of knowledge from a learning environment to real-world settings. Effective use of meta-cognitive skills in the workplace results in more productive employees. Employers will benefit from staff possessed of well developed problem-solving skills. The ability of individuals to resolve ill-defined problems and effectively address open-ended scenarios will benefit society as a whole.

   b. State type, amount, method of disbursement, schedule of payment to be offered, and the effect of withdrawal from participation in the study, if any:

   No payment or disbursement will be offered or provided.

Attachment A
Claim for Exemption
Page 2 of 4
No withdrawal from treatment is expected.

c. Estimated costs to each subject due only to the research participation:

(1). Time (i.e., total time commitment for the duration of the project):

Participants are not expected to spend more or less time than would normally be required for preparation and attendance of a 4-credit hour, undergraduate university course.

(2). Money:

No cost to participants.

(3). Is repeated testing required? Explain:

This study will utilize a time-series design. This design necessitates that an examination may be administered more than once to determine changes in student attitudes and performance. A content-based pre-test will be administered at the beginning of the semester. During the first half of the semester, students will use an Earth-science simulation focusing on concepts from geology. Open completion of the simulation students will complete a content-based exam focusing on concepts introduced through the simulation. A content-based post-test will be administered mid-semester. During the second half of the semester students will use a standardized curriculum without simulations. A content-based post-test will be administered at the end of the semester. The repeated administration of the post-test will provide data on student performance with and without, the inclusion of an instructional simulation into the curriculum.

6) BASIS OF CLAIM FOR EXEMPTION – Federal Regulations and/or University Policy require that in order for research to be exempt from review, at least one of the following blocks letter A-E must be checked.

☐ A. The research will be conducted only in established or commonly accepted educational settings (like classrooms) AND it involves normal educational practices such as research on regular and special educational instructional strategies, or research on the effectiveness of, or the comparison among instructional techniques, curricula or classroom management methods.

☐ B. It will be conducted using only questionnaire or interview survey methods AND the subjects are elected or appointed public officials or candidates for public office.

☐ C. It is limited to the collection and study of existing data, documents, records, pathological or diagnostic specimens which are available to the public.

☐ D. It is limited to the collection and study of obtained using only the following techniques AND the data or information obtained will be recorded in such a manner that subjects cannot be identified, directly or indirectly, through identifiers linked with the subjects.

Check the applicable techniques(s):

☐ The data will be obtained through the use of educational tests (cognitive, diagnostic, aptitude, achievement, etc.), or

☐ Data will be obtained by observing the public behavior of subjects, or

☐ Data will be obtained using survey or interview procedures, or

☐ The data will be obtained from existing documents, records, pathological or diagnostic specimens
E. It is limited to the collection and study of data obtained by
   ☐ observing the public behavior of the participants, or
   ☐ using survey or interview procedures, AND:

   (BOTH OF THE FOLLOWING BOXES MUST BE CHECKED IF THIS BASIS FOR EXEMPTION IS CLAIMED)

   ☑ i.) The information collected about the subjects' behavior DOES NOT INVOLVE sensitive subjects such as illegal or immoral conduct, drug or alcohol abuse, sexual behavior, mental illness, or other possible personally embarrassing subjects, AND

   ☑ ii.) The information collected about subjects, if it became known to outsiders, could not reasonably be expected to place the subject at risk of civil or criminal liability, or be damaging to the subjects social or financial standing or employability.

7) Statement of Risk: The undersigned certify that they believe that the conduct of the above described research creates no more than a minimal risk of physical or emotional harm, or social or legal embarrassment to any participating human subject.

   [Signature]
   8/21/06
   Date

8) Faculty Sponsor (if a student is the Principal Investigator):

   [Signature]
   8/20/06
   Date

9) Recommendation of Program Chair: I recommend that the above described research project be exempt from review.

   [Signature]
   8-29-06
   Date

Attachment A
Claim for Exemption
Page 4 of 4
APPENDIX D

MIDWESTERN STATE UNIVERSITY (MSU)

HUMAN SUBJECTS REVIEW COMMITTEE (HSRC)

INFORMED CONSENT FORM
The same informed consent form was provided by Midwestern State University’s (MSU) Human Subjects Research Committee (HSRC) for the pilot study and for the full study. As the forms were identical except for the course number, only the form used for the full study is reproduced here.
ATTACHMENT C
CONSENT FORM

I hereby give my consent for my participation in the project entitled *An investigation of the use of instructional simulations in the classroom as a methodology for promoting transfer, engagement and motivation.*

I understand that the persons responsible for this project are Dr. Bambi Bailey & Mr. Les M. Lunce, telephone numbers 940-397-6237, 940-766-0485. Dr. Bambi Bailey & Mr. Les M. Lunce have explained that these studies are part of a project that has the following objectives: The goal of this study is to investigate the impact instructional simulations have on the learning experience.

Dr. Bailey’s or Mr. Lunce’s authorized representative has (1) explained the procedures to be followed and identified those that are experimental, (2) described the attendant discomforts and risks, and (3) described the appropriate alternative procedures.

Description of the Project

A content-based pre-test and post-test covering concepts from physics will be administered at the beginning of the Spring 2007 semester. Later in the semester a second content-based pre-test will be administered after which students will use a simulation presenting concepts from physics. Upon completion of the simulation students will complete a post-test focusing on
concepts introduced through the simulation. During the second half of the semester students will use a standardized curriculum without simulations. A third pre-test and post-test covering topics from physics will be administered during the second half of the semester. A fourth pre-test and post-test covering topics from chemistry will be administered in the closing weeks of the semester. Data will be collected from exam scores, attendance records and from a brief survey. Students will be asked to complete the survey at the end of the Spring 2007 semester.

Information concerning payment for my participation in this study has been explained to me as follows: Students will receive no monetary reimbursement or extra credit for participating in this study.

The risks have been explained to me as follows: No risk to participants

It has further been explained to me that the total duration of my participation will be the Spring 2007 semester, that only Dr. Bailey and Mr. Lunce will have access to the records and/or data collected for this study, and that all data associated with this study will remain strictly confidential.

Dr. Bailey & Mr. Lunce have agreed to answer any inquiries I may have concerning the procedures and have informed me that I may contact the Midwestern State University Human Subject Review Committee by writing to: Chair, Human Subjects Review Committee, c/o Office of the Provost, Midwestern State University, 3410 Taft Blvd., Wichita Falls, TX 76308, or by calling the Provost at (940) 397-4226.
If this research project causes any physical injury to participants in this project, treatment is not necessarily available at Midwestern State University Student Health Center, nor is there necessarily any insurance carried by the University or its personnel to cover such an injury. Financial compensation for any such injury must be provided through the participant’s own insurance program. Further information about these matters may be obtained from Chair, Human Subjects Review Committee, c/o Office of the Provost, Midwestern State University, 3410 Taft Blvd., Wichita Falls, TX 76308.

I understand that I may not derive therapeutic treatment from participation in this study. I understand that I may discontinue my participation in this study at any time without penalty.

Signature of Subject __________________________ Date ______________________

Signature of Parent/Guardian or Authorized Representative (if required)

______________________________ Date ______________________

Signature of Witness to Oral Presentation

______________________________ Date ______________________
APPENDIX E

SURVEY
The knowledge-seeking surveys administered in GNSC 1104 for the pilot study and GNSC 1204 for the complete study were both created with the SurveyMonkey® online survey creation software (Finley, R., SurveyMonkey.com, Portland, OR, www.surveymonkey.com). Because of the method used to capture the surveys the page numbers in the lower right corner of each page do not reflect the actual survey screen numbers.
GNSC 1104 Survey, Fall 2006

1. Informed Consent

STATEMENT OF INFORMED CONSENT

Project title: An investigation of the use of instructional simulations in the classroom as a methodology for promoting transfer, engagement and motivation.

I understand that the persons responsible for this project are Dr. Bambi Bailey and Mr. Les M. Lunce, telephone numbers 940-397-6237 and 940-766-0485. Dr. Bailey and/or Mr. Lunce have explained that these studies are part of a project with the objective of investigating the impact of instructional simulations on the learning experience.

Dr. Bailey, Mr. Lunce or their authorized representative has (1) explained the procedures to be followed and identified those that are experimental, (2) described the attendant discomforts and risks, and (3) described the appropriate alternative procedures.

It has been explained to me that I will receive no monetary reimbursement or extra credit for participating in this study.

I understand that the study poses no risk to participants.

It has further been explained to me that the total duration of my participation will be the Fall 2006 semester, that only Dr. Bailey and Mr. Lunce will have access to the records and/or data collected for this study, and that all data associated with this study will remain strictly confidential.

Dr. Bailey & Mr. Lunce have agreed to answer any inquiries I may have concerning the procedures and have informed me that I may contact the Midwestern State University Human Subject Review Committee by writing to: Chair, Human Subjects Review Committee, c/o Office of the Provost, Midwestern State University, 3410 Taft Blvd., Wichita Falls, TX 76308, or by calling the Provost at (940) 397-4262.

If this research project causes any physical injury to participants in this project, treatment is not necessarily available at Midwestern State University Student Health Center, nor is there necessarily any insurance carried by the University or its personnel to cover such an injury. Financial compensation for any such injury must be provided through the participant's own insurance program. Further information about these matters may be obtained from Chair, Human Subjects Review Committee, c/o Office of the Provost, Midwestern State University, 3410 Taft Blvd., Wichita Falls, TX 76308.

I understand that I may not derive therapeutic treatment from participation in this study. I understand that I may discontinue my participation at any time without penalty.
* 1. I have read and understand the terms of informed consent.
   - [ ] I give my consent to participate in this study.
   - [ ] I do not give my consent to participate in this study.

Next >>
2. Identification

* 2. Please enter your social security number or your MSU Campus-Wide ID (CWID).
   Please DO NOT enter dashes (-) or spaces.

* 3. Please indicate which identification number you provided for question 2 above.
   - Social Security Number
   - MSU Campus-Wide ID number

* 4. Have you previously taken one or more courses in Astronomy, Biology, Chemistry, Geology, Physics or any other sciences at a university or college?
   - Yes
   - No

* 5. Where did you work with the Earthquake simulation?
   - Computer lab
   - Home
   - Other
GNSC 1104 Survey, Fall 2006

3. Your opinion of the simulation.

* 6. In your own words, please describe what you learned from the simulation.

* 7. In your own words, please describe what you liked or did not like about using the simulation.
4. Knowledge Seeking 1 of 2

* 0. During the FIRST HALF of the semester while studying earthquakes, did you engage in any of the following activities on your own time, outside of class, to learn more about earthquakes? Please check all activities that apply.

- Watched a prerecorded science program on VHS or DVD about earthquakes.
- Watched a science documentary on The Discovery Channel, The Science Channel, The National Geographic Channel, public television, etc. about earthquakes.
- Watched a satellite, cable or network news program about earthquakes.
- Read all or part of a book on earthquakes.
- Read a magazine or newspaper article about earthquakes.
- Visited the MSU Library or the Public Library to find information about earthquakes.
- Visited a museum or planetarium to see an earthquake exhibit.
- Spoke with your professor, or another professor, about earthquakes.
- Attended a lecture or presentation outside of your regular class activities about earthquakes.
- Searched the Internet for information on earthquakes.
- Other.
- None of the above.
GNSC 1104 Survey, Fall 2006

5. Knowledge Seeking 2 of 2

* 9. During the SECOND HALF of the semester while studying ecology, did you engage in any of the following activities on your own time, outside of class, to learn more about Biological Science? Please check all activities that apply.

- Watched a prerecorded science program on VHS or DVD on ecology.
- Watched a science documentary on The Discovery Channel, The Science Channel, The National Geographic Channel, public television, etc.
- Watched a satellite, cable or network news program which covered ecology.
- Read all or part of a book on ecology.
- Read a magazine or newspaper article on ecology.
- Visited the MSU Library or the Public Library to find information about ecology.
- Visited a museum or planetarium to see an exhibit on ecology.
- Spoke with your professor, or another professor, about ecology.
- Attended a lecture or presentation outside of your regular class activities on ecology.
- Searched the Internet for information on ecology.
- Other.
- None of the above.
6. Thanks!

Thank you for taking time to complete this survey.
Welcome to Midwestern State University's Gordon T. and Ellen West College of Education

A Welcome from the Dean:

Welcome to the Gordon T. and Ellen West College of Education, one of the best small colleges of education. The West College offers undergraduate teacher education programs for elementary and secondary teachers as well as post-baccalaureate certification for persons interested in a career in teaching who already have a baccalaureate degree. The college offers baccalaureate degrees in curriculum and instruction, educational leadership, reading, special education, school counseling, and master of Arts degrees in counseling, human resource development, and training and development.

The West College of Education has an approved Candidate for the Professional Development of Teachers with seven professional development schools located in area public schools. The undergraduate teacher education program is field based with all professional development courses taught in the professional development schools. As a result, each professional education course integrates theory and practice. Our students, therefore, enter the teaching profession with considerable experience and confidence. This integration of theory and practice combines with the alignment of our curriculum with the Texas standards for teachers has resulted in our students having one of the highest success rates on the state required teacher certification examinations (TECEX). In 1998-1999, more than 90% of our students passed all required EXCE exam during their first year.

In the past sixteen years the West College of Education has received more than $3 million in grants from the West Foundation of Wichita Falls. This funding has enabled the college to attract outstanding faculty and has provided us with the highest quality teaching technology. Our small classes, close relationship with teaching professionals, the support of both the West Foundation, and the university administration result in a joyful atmosphere for teaching and learning.

Grant W. Simpson, Ph. D.
Dean, West College of Education
GNSC 1104 Survey, Fall 2006

1. Informed Consent

STATEMENT OF INFORMED CONSENT

Project title: An investigation of the use of instructional simulations in the classroom as a methodology for promoting transfer, engagement and motivation.

I understand that the persons responsible for this project are Dr. Bambi Bailey and Mr. Les M. Lunce, telephone numbers 940-397-6237 and 940-766-0485. Dr. Bailey and/or Mr. Lunce have explained that these studies are part of a project with the objective of investigating the impact of instructional simulations on the learning experience.

Dr. Bailey, Mr. Lunce or their authorized representative has (1) explained the procedures to be followed and identified those that are experimental, (2) described the attendant discomforts and risks, and (3) described the appropriate alternative procedures.

It has been explained to me that I will receive no monetary reimbursement or extra credit for participating in this study.

I understand that the study poses no risk to participants.

It has further been explained to me that the total duration of my participation will be the Fall 2006 semester, that only Dr. Bailey and Mr. Lunce will have access to the records and/or data collected for this study, and that all data associated with this study will remain strictly confidential.

Dr. Bailey & Mr. Lunce have agreed to answer any inquiries I may have concerning the procedures and have informed me that I may contact the Midwestern State University Human Subject Review Committee by writing to: Chair, Human Subjects Review Committee, c/o Office of the Provost, Midwestern State University, 3410 Taft Blvd., Wichita Falls, TX 76308, or by calling the Provost at (940) 397-4226.

If this research project causes any physical injury to participants in this project, treatment is not necessarily available at Midwestern State University Student Health Center, nor is there necessarily any insurance carried by the University or its personnel to cover such an injury. Financial compensation for any such injury must be provided through the participant’s own insurance program. Further information about these matters may be obtained from Chair, Human Subjects Review Committee, c/o Office of the Provost, Midwestern State University, 3410 Taft Blvd., Wichita Falls, TX 76308.

I understand that I may not derive therapeutic treatment from participation in this study. I understand that I may discontinue my participation at any time without penalty.
* 1. I have read and understand the terms of informed consent.
   - [ ] I give my consent to participate in this study.
   - [ ] I do not give my consent to participate in this study.

   Next >>
2. Identification

* 2. Please enter your social security number or your MSU Campus-Wide ID (CWID). Please DO NOT enter dashes (-) or spaces.

* 3. Please indicate which identification number you provided for question 2 above.
   - Social Security Number
   - MSU Campus-Wide ID number

* 4. Have you previously taken one or more courses in Astronomy, Biology, Chemistry, Geology, Physics or any other sciences at a university or college?
   - Yes
   - No

* 5. Where did you work with the sound wave simulation?
   - Computer lab
   - Home
   - Other

<< Previous    Next >>
3. Learning experience with a simulation.

* 6. In your own words, please describe what you learned from the simulation.

* 7. In your own words, please describe how you learned differently from the simulation than you have learned from other experiences (classroom lecture, lab, etc.)?

* 8. In your own words, please describe the effect the simulation had on your learning.

* 9. In your own words, please describe what you liked or did not like about using the simulation.
4. Transfer of knowledge and skills.

* 10. In your own words, please describe the effect using the simulation had on the way you went about studying electricity and the periodic table.
## GNSC 1204 Survey, Spring 2007

### 5. Knowledge Seeking

**11. Did you engage in any of the following activities on your own time, outside of class, to learn more about Physics and Chemistry? Please check all activities that apply.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Watched a prerecorded science program on VHS or DVD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watched a science documentary on The Discovery Channel, The Science Channel, The National Geographic Channel, public television, etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watched a satellite, cable or network news program.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read all or part of a book.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read a magazine or newspaper article.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visited the MSU Library or the Public Library.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visited a museum or planetarium.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spoke with your professor, or another professor.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attended a lecture or presentation outside of your regular class activities.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Searched the Internet.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other activity(s) not shown.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None of the above.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
GNSC 1204 Survey, Spring 2007

6. Thanks!

Thank you for taking time to complete this survey.

<< Previous       Done >>
Welcome to Midwestern State University's Gordon T. and Ellen West College of Education

A Welcome from the Dean:

Welcome to the Gordon T. and Ellen West College of Education, one of the best small colleges of education. The West College offers undergraduate teacher education programs for elementary and secondary teachers as well as post-baccalaureate certification for persons interested in a career in teaching who already have a baccalaureate degree. The college offers a B.S. in Liberal Studies, an education degree in curriculum and instruction, educational leadership, reading, special education, school counseling, and master of Arts degrees in counseling, human resource development, and training and development.

The West College of Education has an approved program for the Professional Development of Teachers within the professional development schools located in area public schools. The undergraduate teacher education program is field based with all professional development courses taught in the professional development schools. As a result, each professional education course integrates theory and practice. Our students, therefore, enter the teaching profession with considerable experience and confidence. This integration of theory and practice combined with the alignment of our curriculum with the Texas standards for teachers has resulted in our students having one of the highest success rates on the state required teacher certification examinations (ACT). In 1999-2000, more than 90% of our students passed all required EXCET tests during their first year.

In the past sixteen years the West College of Education has received more than $3 million in grants from the West Foundation of Wichita Falls. This funding has enabled the college to attract outstanding faculty and has provided us with the highest quality teaching technology. Our small classes, close relationship with teaching candidates, the support of both the West Foundation, and the university administration result in a joyful atmosphere for teaching and learning.

Grant W. Simpson, Ph. D.
Dean, West College of Education
APPENDIX F

ACT-SAT SCORE CONVERSION TABLE
Table 1

**ACT-SAT Score Comparisons**

<table>
<thead>
<tr>
<th>SAT Score Verbal + Math</th>
<th>ACT Composite Score</th>
<th>ACT Composite Score</th>
<th>SAT Score Verbal + Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>1600</td>
<td>36</td>
<td>36</td>
<td>1600</td>
</tr>
<tr>
<td>1560–1590</td>
<td>35</td>
<td>35</td>
<td>1580</td>
</tr>
<tr>
<td>1510–1550</td>
<td>34</td>
<td>34</td>
<td>1520</td>
</tr>
<tr>
<td>1460–1500</td>
<td>33</td>
<td>33</td>
<td>1470</td>
</tr>
<tr>
<td>1410–1450</td>
<td>32</td>
<td>32</td>
<td>1420</td>
</tr>
<tr>
<td>1360–1400</td>
<td>31</td>
<td>31</td>
<td>1380</td>
</tr>
<tr>
<td>1320–1350</td>
<td>30</td>
<td>30</td>
<td>1340</td>
</tr>
<tr>
<td>1280–1310</td>
<td>29</td>
<td>29</td>
<td>1300</td>
</tr>
<tr>
<td>1240–1270</td>
<td>28</td>
<td>28</td>
<td>1260</td>
</tr>
<tr>
<td>1210–1230</td>
<td>27</td>
<td>27</td>
<td>1220</td>
</tr>
<tr>
<td>1170–1200</td>
<td>26</td>
<td>26</td>
<td>1180</td>
</tr>
<tr>
<td>1130–1160</td>
<td>25</td>
<td>25</td>
<td>1140</td>
</tr>
<tr>
<td>1090–1120</td>
<td>24</td>
<td>24</td>
<td>1110</td>
</tr>
<tr>
<td>1060–1080</td>
<td>23</td>
<td>23</td>
<td>1070</td>
</tr>
<tr>
<td>1020–1050</td>
<td>22</td>
<td>22</td>
<td>1030</td>
</tr>
</tbody>
</table>

*(table continued)*
Table 1 (continued).

<table>
<thead>
<tr>
<th>SAT Score Verbal + Math</th>
<th>ACT Composite Score</th>
<th>ACT Composite Score</th>
<th>SAT Score Verbal + Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>980–1010</td>
<td>21</td>
<td>21</td>
<td>990</td>
</tr>
<tr>
<td>940–970</td>
<td>20</td>
<td>20</td>
<td>950</td>
</tr>
<tr>
<td>900–930</td>
<td>19</td>
<td>19</td>
<td>910</td>
</tr>
<tr>
<td>860–890</td>
<td>18</td>
<td>18</td>
<td>870</td>
</tr>
<tr>
<td>810–850</td>
<td>17</td>
<td>17</td>
<td>830</td>
</tr>
<tr>
<td>760–800</td>
<td>16</td>
<td>16</td>
<td>780</td>
</tr>
<tr>
<td>710–750</td>
<td>15</td>
<td>15</td>
<td>740</td>
</tr>
<tr>
<td>660–700</td>
<td>14</td>
<td>14</td>
<td>680</td>
</tr>
<tr>
<td>590–650</td>
<td>13</td>
<td>13</td>
<td>620</td>
</tr>
<tr>
<td>520–580</td>
<td>12</td>
<td>12</td>
<td>560</td>
</tr>
<tr>
<td>500–510</td>
<td>11</td>
<td>11</td>
<td>500</td>
</tr>
</tbody>
</table>

*Note. As documented in Dorans, (1999), Dorans, Lyu, Pommerich and Houston, (1997) and, Schneider and Dorans, (1999), “equivalent scores are those with the same percentile ranks for a common group of test-takers. A concordance table is dependent upon the sample used to establish the relationship between two sets of scores. Data in this table are based on 103,525 test-takers who took both the SAT and the ACT between October 1994 and December 1996. SAT scores do not cover the full range of the ACT scale due to differences in how percentiles are distributed at the top and bottom of the two scales.”*
APPENDIX G

MIDWESTERN STATE UNIVERSITY (MSU)

STUDENT DEMOGRAPHIC DATA
Table 2

*MSU Enrollment Data, Fall 2006*

<table>
<thead>
<tr>
<th></th>
<th>Students in GNSC 1104</th>
<th>Students in West COE</th>
<th>All Undergrads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age</td>
<td>21</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>Males</td>
<td>10.3%</td>
<td>31.0%</td>
<td>42.3%</td>
</tr>
<tr>
<td>Females</td>
<td>89.7%</td>
<td>69.0%</td>
<td>57.7%</td>
</tr>
<tr>
<td>White</td>
<td>72.0%</td>
<td>80.5%</td>
<td>66.8%</td>
</tr>
<tr>
<td>African American</td>
<td>18.0%</td>
<td>7.5%</td>
<td>13.2%</td>
</tr>
<tr>
<td>Hispanic American</td>
<td>6.0%</td>
<td>8.3%</td>
<td>9.3%</td>
</tr>
<tr>
<td>Asian American</td>
<td>1.5%</td>
<td>1.0%</td>
<td>3.7%</td>
</tr>
<tr>
<td>Native American</td>
<td>0.0%</td>
<td>1.3%</td>
<td>0.9%</td>
</tr>
<tr>
<td>Non-resident alien</td>
<td>0.0%</td>
<td>1.3%</td>
<td>5.2%</td>
</tr>
<tr>
<td>Not reported</td>
<td>1.5%</td>
<td>0.3%</td>
<td>0.9%</td>
</tr>
</tbody>
</table>

*Note.* †Institutional Research, MSU, Wichita Falls, Texas, spring 2007.
Table 3

*MSU Enrollment Data, Spring 2007*

<table>
<thead>
<tr>
<th></th>
<th>Students in GNSC 1204</th>
<th>Students in West COE</th>
<th>All MSU Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age</td>
<td>23</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>Males</td>
<td>8.0%</td>
<td>9.0%</td>
<td>42.0%</td>
</tr>
<tr>
<td>Females</td>
<td>92.0%</td>
<td>91.0%</td>
<td>58.0%</td>
</tr>
<tr>
<td>White</td>
<td>79.5%</td>
<td>79.8%</td>
<td>66.7%</td>
</tr>
<tr>
<td>African American</td>
<td>7.7%</td>
<td>7.5%</td>
<td>13.6%</td>
</tr>
<tr>
<td>Hispanic American</td>
<td>9.0%</td>
<td>9.0%</td>
<td>8.8%</td>
</tr>
<tr>
<td>Asian American</td>
<td>1.3%</td>
<td>1.2%</td>
<td>4.1%</td>
</tr>
<tr>
<td>Native American</td>
<td>0.0%</td>
<td>0.9%</td>
<td>0.8%</td>
</tr>
<tr>
<td>Non-resident alien</td>
<td>2.6%</td>
<td>2.4%</td>
<td>5.1%</td>
</tr>
<tr>
<td>Not reported</td>
<td>0.0%</td>
<td>0.3%</td>
<td>0.9%</td>
</tr>
</tbody>
</table>

*Note.* Institutional Research, MSU, Wichita Falls, Texas, spring 2007.
Table 4

*Student Population Data, Texas*

<table>
<thead>
<tr>
<th></th>
<th>MSU&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Angelo State&lt;sup&gt;2&lt;/sup&gt;</th>
<th>West Texas A&amp;M&lt;sup&gt;3&lt;/sup&gt;</th>
<th>UT Tyler&lt;sup&gt;4&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age</td>
<td>22</td>
<td>18</td>
<td>26</td>
<td>21</td>
</tr>
<tr>
<td>Males</td>
<td>42.0%</td>
<td>46.0%</td>
<td>42.0%</td>
<td>40.0%</td>
</tr>
<tr>
<td>Females</td>
<td>58.0%</td>
<td>54.0%</td>
<td>58.0%</td>
<td>60.0%</td>
</tr>
<tr>
<td>White</td>
<td>66.7%</td>
<td>68.5%</td>
<td>75.0%</td>
<td>78.8%</td>
</tr>
<tr>
<td>African American</td>
<td>13.6%</td>
<td>6.0%</td>
<td>3.9%</td>
<td>9.8%</td>
</tr>
<tr>
<td>Hispanic American</td>
<td>8.8%</td>
<td>22.6%</td>
<td>15.0%</td>
<td>5.3%</td>
</tr>
<tr>
<td>Asian American</td>
<td>4.1%</td>
<td>1.3%</td>
<td>1.4%</td>
<td>1.8%</td>
</tr>
<tr>
<td>Native American</td>
<td>0.8%</td>
<td>0.3%</td>
<td>1.0%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Non-resident alien</td>
<td>5.1%</td>
<td>1.2%</td>
<td>2.0%</td>
<td>1.4%</td>
</tr>
<tr>
<td>Unknown</td>
<td>0.9%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>1.9%</td>
</tr>
</tbody>
</table>

Table 5

*Student Population Data, Outside Texas*

| 2005-2006 Academic Year, MSU Student Population Compared to Public Universities Outside Texas*8 |
|------------------------------------------|----------|------------------|------------------|------------------|
| MSU1 | Nicholls State University5 | University of California Stanislaus6 | Clarion University of Pennsylvania7 |
| Mean age  | 22  | 22  | 22  | 22  |
| Males  | 42.0% | 36.75 | 33.0% | 39.0% |
| Females  | 58.0% | 63.3% | 66.0% | 60.1% |
| White  | 66.7% | 74.2% | 41.0% | 91.0% |
| African American  | 13.6% | 18.5% | 3.65% | 5.0% |
| Hispanic American  | 8.8% | 1.8% | 28.0% | 0.8% |
| Asian American  | 4.1% | 1.3% | 11.62% | 0.7% |
| Native American  | 0.8% | 2.0% | 0.9% | 0.2% |
| Non-resident alien  | 5.1% | n/a | n/a | 0.6% |
| Unknown  | 0.9% | 2.2% | 13.0% | 0.9% |

A pilot study was conducted during the fall 2006 semester to test data collection methods and to correct procedural flaws (Gall, Gall and Borg, 2003; Isaac and Michael, 1997). Anticipated benefits to be obtained from conducting the pilot study are outlined in Gall et al. (2003) and Isaac and Michael (1997). First, the pilot study provided an opportunity to test the hypothesis on which this the principle study was based. Second, the pilot study yielded insights with regard to optimal data collection and analysis methods. Third, the pilot study provided an opportunity to test and evaluate the correctness of statistical procedures to be used. Fourth, treatment errors may have been significantly reduced as a result of insights gained from the pilot study. Finally, feedback obtained from participants and others involved in the pilot study may help to improve research methods and indicate unforeseen validity threats.

The population for the pilot study consisted of 5,544 full-time undergraduate students enrolled at MSU during the 2006 fall term. Midwestern State University (MSU) is an independent, public university situated in the North-central Texas city of Wichita Falls (Region 9). The student population demographic of MSU for the fall 2006 term was composed of 57% women, 43% men, 13% African American, 3% Asian American or Pacific Islander, 9% Hispanic American, 1% Native American, international population 5% representing 37 other countries.

The pilot study focused on one section of GNSC 1104 Life/Earth Science, an undergraduate course offered by the Department of Mathematics and Science at MSU. The course was taught by an Associate Professor of Education and Reading and is part of the core curriculum for all undergraduates in the West College of Education at MSU. Course enrollment provided a convenience sample of \( n = 68 \).
A non-randomized control-group pretest-posttest research design was used (Isaac and Michael, 1997). This design is documented in detail in Shadish, Cook and Campbell (2002) as the untreated control group design with dependent pretest and posttest samples. See also Cook and Campbell (1979)

\[ O_{A1} \times O_{A2} \]
\[ O_{B1} \quad O_{B2} \]

This design has demonstrated good external validity because it utilizes intact groups (Isaac and Michael, 1997). Threats from reactive effects of experimental procedures are also minimized by this design. Internal validity of the pilot study was strengthened by standardizing all scores to z-scores. The presence of the control helped to diminish treats from history, maturation and instrumentation. Mortality effects were controlled by discarding any observations with missing test score values. The model also somewhat reduced threats arising from statistical regression to the mean. Internal validity threats which could not be extensively controlled for were selection and maturation interaction, selection and history and, selection and testing (Campbell and Stanley, 1969; Isaac and Michael, 1997).

Scores were collected from content-specific pretests and posttests administered by the Associate Professor. A pretest was administered during the first half of the semester to measure student’s pre-existing knowledge of Earth science. Students then completed a unit on geology in which the Virtual Earthquake© <http://www.sciencecourseware.org/eec/Earthquake/> simulation was used as the dominant curriculum (Deshamais et al., 2002). A posttest was administered after students completed the simulation. The unit incorporated three chapters of course
material with curriculum activities centered on the Virtual Earthquake© simulation. This simulation constituted the treatment phase of the pilot study.

The Virtual Earthquake© simulation was developed as part of the Virtual Courseware Project©, a project supported and validated by the U.S. National Science Foundation and the California State University System. Authors of the Virtual Courseware Project© are Robert A Desharnais, Ph.D., Department of Biology and Microbiology, California State University, Los Angeles, Gary Novak, Ph.D., Department of Geology, California State University, Los Angeles, and David Mayo, Ph.D., Department of Geology, California State University, Los Angeles. Additional validation of the simulation software has been provided by The Center of Usability in Design and Assessment (CUDA) of California State University, Long Beach, (http://www.csulb.edu/centers/cuda/). CUDA conducted a formal teaching effectiveness evaluation of the Virtual Earthquake© simulation (Novak, 1999). Results were published in a 29-page report prepared by Fisk et al., (1998). Data for the assessment were collected from 38 instructors in 14 countries who used the simulation as part of their regular course curriculum. At least 75% of the instructors surveyed indicated that the Virtual Earthquake© simulation had a positive impact on learning effectiveness, speed of learning and student motivation. Negative comments regarding the simulation focused on the poor quality of some graphics and the necessity for repetitive scrolling on some screens.

During the second half of the fall 2006 semester a second content-specific pretest was administered to gage existing student knowledge of biological sciences. Students then used a textbook-based curriculum to complete a unit on ecology. A
content-specific posttest was administered at the end of the unit. Once again, a unit included three chapters of course content. This part of the curriculum represented the control component of the pilot study. Students were also required to complete a lab exercise and daily quizzes through WebCT©. These scores were collected as documentation of attendance and as indicators of academic performance throughout the semester. At the end of term students were asked to complete a brief survey. As part of the survey students were asked to self-report on knowledge-seeking activities in which they engaged during the semester. All data were collated and analyzed using Statistical Package for the Social Sciences (SPSS).

Data

The pilot study setting provided a convenience sample of $n = 71$ observations. Of these observations 3 were discarded because more then two test scores were missing. This resulted in a sample size of $n = 68$.

To increase readability of the data, the gender variable was recoded from string to numeric with the values changed from F to 0 for female and M to 1 for male (Leech et al., 2005).

Three values (4%) for the lab (lab exercise grade) variable were missing. These missing values were replaced with the mean value for the variable (Leech et al., 2005; Montgomery, 1984).

Both the variables sat (SAT score) and act (ACT score) had a number of missing values. When an ACT score / SAT score was present without an SAT score / ACT score, the corresponding SAT score / ACT score was converted to ACT score / SAT score and inserted into the dataset to replace the missing value. Replacement scores
were taken from a standardized ACT-SAT conversion table. See Appendix G. When both ACT and SAT scores were not present, the mean score was used to replace the missing SAT score (Leech et al., 2005; Montgomery, 1984). The missing ACT score was obtained from a standard conversion table. See Appendix G. This replacement occurred 14 times (21% of observations).

For the variable hsgpa (high school cumulative grade point average), 33 (49%) values were missing from the dataset. Rather than attempting to use a replacement method to estimate missing values resulting in potentially serious validly issues, this variable was not used in the pilot study analysis (Leech et al., 2005).

The variables testa1 (pretest A1), testa2 (posttest A2), testb1 (pretest B1, testb2 (posttest B2), lab (laboratory exercise grade), quiz (composite quiz grade) and crsgrade (course grade) were all measured on different scales. To obtain meaningful results it was necessary to perform a simple transformation to place all scores on the same scale. The transformation was performed in two steps. First, all scores were converted to standard (z) scores (Diekhoff, 1992).

$$Z_i = \frac{(X_i - \overline{X})}{SD_X}$$

Second, z-scores were transformed using the following formula with standard deviation (SD) = 10 and the sample mean = 50 (Cronk, 2004; Leech et al., 2005).

$$t = SD \times z + \text{mean}$$

The large sample size ($n = 68$) strengthened reliability and internal validity of the pilot study results, reduced the probability of errors arising from sampling, and increased the power of statistical tests used (Isaac and Michael, 1997). Sampling error was estimated using the following formula (Diekhoff, 1992).
Margin of error (95%) = 1.96 × \sqrt{\frac{0.5(1 - 0.5)}{n}} = \frac{0.98}{\sqrt{n}}

This formula yielded a sampling error estimate (e) of ±0.119 given a confidence interval
Z = 1.96 (95%), (p ≤ .05).

Interpretation of Results

Results obtained from descriptive and frequency statistics were mixed and therefore inconclusive. Distributions for all variables measured exhibited some degree of skewness and kurtosis (Gall et al., 2003; Hinkle et al., 2003; Isaac and Michael, 1997). However, the degree of skewness and kurtosis were deemed insufficient to risk validity threats which could arise from applying a transformation (Gall et al., 2003; Hinkle et al., 2003; Isaac and Michael, 1997; Leech et al., 2005; Montgomery, 1984). The sample was of sufficient size (n = 68) and the distributions were close enough to normal that threats to homogeneity of variance were minimized.

The homogeneity of variance assumption concerns population variances (Hinkle et al., 2003). The impact of violating this assumption depends on the size of the samples and can significantly impact results obtained from t-test and analysis of variance (ANOVA) (Leech, et al., 2005). If the sample is of sufficient size, then the effect of violating the assumption is not serious (Hinkle et al., 2003). The one-way ANOVA test analyzes a single variable with more than one level by looking for variance among the level means (Hinkle, et al., 2003). A one-way ANOVA was computed using six predictor variables (testa1, testa2, testb1, testb2, lab, quiz) to check the assumption of homogeneity of variance. Test results are listed in Table 6.
Levene's test was computed to test the hypothesis that the variances among the six sets of scores were different. The analysis of variance (ANOVA) value of \( \text{Sig.} \) 0.253 was > .05 and was not significant. The hypothesis that the variances among the six groups of scores were significantly different was rejected. If \( \text{Sig.} \) were < .05 it would have indicated that the assumptions of ANOVA had been violated and a transformation would be warranted. As the homogeneity of variance assumption was met no further transformations were computed (Muijs., 2004).

Covariance is defined as the average cross-product of deviation scores, or the average cross-product of two score’s deviation from the mean (Gall et al., 2003; Hinkle et al., 2003; Isaac and Michael, 1997).

\[
COV_{XY} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{n - 1}
\]

Covariance indicates whether two scores are correlated or deviate together, whether there is a relationship between two scores, and whether that relationship is positive or negative. If the covariance is 0, no relationship can be assumed to exist between the two scores. If covariance is positive then both variables change in the same direction. If covariance is negative then when one variable changes, the other changes in the opposite direction.
Covariance among the six predictor variables (testa1, testa2, testb1, testb2, lab, quiz) indicated positive relationships. However, the covariance was high for only two pairs of variables, testa1 and testb1 (COV = 102.03, r = .647 moderate) and testa1 and lab (COV = 254.13, r = .336 low). These two pairs of variables exhibited a strong positive relationship. Table 7 lists covariance among the 6 predictor variables.

Table 7

<table>
<thead>
<tr>
<th></th>
<th>testa2</th>
<th>testb1</th>
<th>testb2</th>
<th>quiz</th>
<th>lab</th>
</tr>
</thead>
<tbody>
<tr>
<td>COV</td>
<td>99.54</td>
<td>102.03</td>
<td>6.38</td>
<td>63.79</td>
<td>254.13</td>
</tr>
<tr>
<td>r</td>
<td>.927</td>
<td>.647</td>
<td>.619</td>
<td>.597</td>
<td>.336</td>
</tr>
<tr>
<td>COV</td>
<td>63.84</td>
<td>74.93</td>
<td>14.09</td>
<td>15.58</td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>.594</td>
<td>.654</td>
<td>.782</td>
<td>.654</td>
<td></td>
</tr>
<tr>
<td></td>
<td>63.79</td>
<td>15.58</td>
<td>.143</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>.597</td>
<td>.358</td>
<td>.158</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>254.13</td>
<td>15.58</td>
<td>.158</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>.336</td>
<td>.358</td>
<td>.158</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The paired-samples or dependent-samples *t-test* compares the means of two scores from related samples, i.e. pretest and posttest scores from a single sample (Cronk, 2004; Gall et al., 2003; Hinkle et al., 2003). The hypothesis that differences exist between mean examination scores is tested by measuring the same individual on the same construct, on two different occasions (Gall et al., 2003; Hinkle et al., 2003). Both variables must be interval or ratio scale and scores must be normally distributed. If the scales are different then a conversion to z-scores is required (Gall et al., 2003; Hinkle et al., 2003).
A paired-samples \textit{t-test} was computed to investigate correlations between scores. With $n = 68$, degrees of freedom (\textit{df}) = 67, $\alpha = .05$, $p < .001$, values for \textit{t-calculated} ($t$) were obtained from the paired samples test table generated by SPSS. The value for \textit{t-critical} = 2.660 was obtained from the critical values of the t-distribution table (Hinkle et al., 2003). For all variables, \textit{t-calculated} fell within the confidence interval (CI), and \textit{t-calculated} was $< \text{t-critical}$. Values for \textit{t-calculated} did not indicate any significant differences between scores. The hypotheses that means were significantly different was not supported. Results were not statistically significant.

For all pairs of variables, \textit{p-calculated} (Sig. 2-tailed) was $> \text{p-critical} = .001$ and/or \textit{p-calculated} fell within the CI. Once again, the hypothesis that means were significantly different was not supported. Results were not statistically significant.

When data are close to the mean, not spread out, the probability of observing statistically significant results, rejecting the null hypotheses, is increased. In the case of this pilot study, scores were generally close to the mean. However, the more spread out the means, the greater the difference between the means, the greater the likelihood that statistically significant results will be observed. In this pilot study, with one exception, means were close together so observation of statistically significant results was not anticipated.

The closer the standard deviations, the less difference between the standard deviations, the greater the likelihood that statistically significant results will be observed. Standard deviations in this pilot study were spread out, 4.0914 to 13.0913 which precluded the observation statistically significant results.
Multiple linear regression (MLR) tests the hypothesis that values for one variable, the criterion or dependent variable, can be predicted based on the values of one or more independent or predictor variables (Cronk, 2004; Gall et al., 2003; Hinkle et al., 2003; Leech et al., 2005). An MLR was calculated to predict course grade from 10 predictor variables. Partial MLR results are listed in Table 8 and Table 9. A significant regression equation was found.

Table 8
ANOVA Model Summary

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th>R</th>
<th>(R^2)</th>
<th>Adjusted (R^2)</th>
<th>Std. Error of Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0.989</td>
<td>0.978</td>
<td>0.974</td>
<td>1.6213</td>
</tr>
</tbody>
</table>

\((F(10,57) = 249.175, p < .001)\), with an \(R^2\) of .978.

\[ Y^\prime = \alpha + \beta_0X_0 + \beta_1X_1 + \beta_2X_2 + \ldots \beta_nX_n \]

\[ Y^\prime = (-7.213) + (0.179*testa1) + (0.017*testa2) + (0.299*testb1) + (0.459*testb2) + (0.114*quiz) + (0.152*lab) + (-0.032*sat) + (-0.041*age) + (-0.003*ethnic) + (0.023*gender). \]

MLR results indicated that 5 of the predictor variables, testa1, testa2, testb1, testb2 and quiz, were significantly correlated with course grade. The other 5 predictors were moderately or weakly correlated with course grade. Two predictor variables contributed significantly to the variance in course grade as indicated by their structure coefficients, \(t\) and \(Sig.\) values; testb2 (.903; 12.118; .000) and testb1 (.817; 7.692; .000). Only testb1, testb2 and lab resulted in a \(Sig.\) value of 0.000 at \(\alpha = 0.05\) level with a \(t\) value for lab of 6.898. The high correlations between testa1 and testa2 as well as between testb1 and testb2 indicated high test-retest reliability and criterion-related
### Table 9

**Correlations with Course Grade**

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Pearson</th>
<th>Sig. (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlations with</td>
<td>Sig. Level of</td>
</tr>
<tr>
<td></td>
<td>Course Grade</td>
<td>Correlations</td>
</tr>
<tr>
<td>testa1</td>
<td>.784</td>
<td>.000</td>
</tr>
<tr>
<td>testa2</td>
<td>.743</td>
<td>.000</td>
</tr>
<tr>
<td>testb1</td>
<td>.890</td>
<td>.000</td>
</tr>
<tr>
<td>testb2</td>
<td>.915</td>
<td>.000</td>
</tr>
<tr>
<td>quiz</td>
<td>.826</td>
<td>.000</td>
</tr>
<tr>
<td>lab</td>
<td>.352</td>
<td>.002</td>
</tr>
<tr>
<td>sat</td>
<td>.355</td>
<td>.001</td>
</tr>
<tr>
<td>age</td>
<td>.010</td>
<td>.459</td>
</tr>
<tr>
<td>ethnicity</td>
<td>-.390</td>
<td>.001</td>
</tr>
<tr>
<td>gender</td>
<td>.099</td>
<td>.210</td>
</tr>
</tbody>
</table>

Validity. The Sig. (p-calc) value indicated the probability of obtaining the corresponding t value. At $\alpha = .05$ significance level, results for testa1, testb1, testb2, quiz and lab were less than .05 and so were significant. Results from other predictors were not significant.

A *synthetic variable* ($Y^\wedge$) was computed for each observation. The sum of $Y^\wedge$ and $e$ equaled $Y$ (course grade) for each observation. These results validated the replicability of the pilot study results.
From the MLR model summary table, residuals yielded a multiple $R$ of 0.989, an effect size ($R^2$) of 0.978 and an adjusted $R^2$ of 0.974. The ANOVA summary table yielded statistically significant results of .000 at $\alpha = 0.05$. These results supported the hypothesis that the required assumptions of multiple regression were met (Gall et al., 2003).

Very little shrinkage between effect size ($R^2$) = 0.978 and adjusted $R^2$ = 0.974 was indicated. This minimal sampling error supported the hypothesis that the sample (n) was closely representative of the population (Gall et al., 2003).

With a coefficient of determination, $R^2 = .978$, 97.8% of the variance in the dependent / criterion variable, course grade, was explained or predicted by variance in the 10 predictor variables. This strong effect size indicated that further examination of the data was warranted. A certain percentage of effect size arose from error ($e$) defined as the coefficient of alienation $= 1 – 0.978 = 0.022$. The error was small so the regression equation was a good predictor of course grade.

The standard error of the estimate from the MLR model summary table, 1.6213, represented a margin of error for the prediction equation. This value was small further supporting a high degree of prediction between the predictor and criterion variables. The model was a strong fit with adjusted $R^2 = .974 > .05$ (Muijs, 2004).

From the MLR coefficients table, $B$ coefficients indicated by how much the value of the variable $Y$ (course grade) changed if the value of the variable $X$ changed by 1 unit. For example, if lab increased by 1, course grade increased by .152.

The standard error of the $B$ coefficient was deemed dependable because all scores were converted to transformed z-scores prior to running the regression. The
largest $\beta$ weight was indicated for testb2. Because the six scores were standardized and transformed prior to running the regression, the *un-standardized B coefficients* and the standardized $\beta$ coefficients were identical.

*Sig.* values from the *MLR* coefficients table indicated a statistically significant relationship between each predictor and the criterion variable, course grade. *Sig.* values represented the probability of finding a specific relationship, of a given magnitude, between a given predictor and the criterion in the sample when that same relationship did not exist in the population. At $p < .001$ only testb1, testb2 and lab exhibit statistically significant relationships with course grade. The $t$ and *Sig.* values from the *MLR* coefficients table also indicated whether each independent variable was contributing significantly to the regression equation for predicting course grade from the set of 10 predictors. Only three predictor variables were observed to contribute significantly to the prediction of course grade as shown in Table 10.

**Table 10**

*Significant Predictor Variables*

<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>Sig. &lt; .001</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>testb2</td>
<td>12.118</td>
<td>.000</td>
<td>.459</td>
</tr>
<tr>
<td>testb1</td>
<td>7.692</td>
<td>.000</td>
<td>.299</td>
</tr>
<tr>
<td>lab</td>
<td>6.898</td>
<td>.000</td>
<td>.152</td>
</tr>
</tbody>
</table>

Based on results from the *MLR ANOVA* summary table, *Sig.* of .001 < .05, the predictor variables significantly predicted the criterion variable. These results were
statistically significant. Given $F_{calculated} = 249.175$ with $df(10, 57)$ from the ANOVA summary table and $\alpha = .05$, from the $F$-distribution table (Hinkle et al., 2003), an $F$-critical value of 2.08 was obtained. $F_{calculated}$ was $> F$-critical ($249.175 > 2.08$). Once again results were statistically significant. Based on values from the ANOVA summary table, $\eta^2 = SS_{Regression} / SS_{Total} = 6550.220 / 6700.059 = 0.978$. This value indicated that 97.8% of the variance in course grade was explained by the variance in the other variables. These results were in agreement with results obtained from the Model Summary table.

Repeated measures ANOVA tests the hypothesis of both between-groups factors and within-subjects (repeated-measures) factors (Cronk, 2004; Gall et al., 2003; Hinkle et al., 2003; Leech et al., 2005). Interaction between factors as well as effects of individual factors on one or more dependent variables can be investigated (Gall et al., 2003; Hinkle et al., 2003). Repeated measures ANOVA requires that study participants provide data on more then one level of an independent variable (Gall et al., 2003; Hinkle et al., 2003). Partial results of the repeated measures ANOVA test computed for this pilot study are shown in Table 11, Table 12, and Table 13.

**Table 11**

*Repeated Measures ANOVA*

<table>
<thead>
<tr>
<th>Within Subjects Effect</th>
<th>Mauchly’s W</th>
<th>Approx. Chi-Square</th>
<th>df</th>
<th>Sig</th>
<th>Greenhouse-Geisser</th>
<th>Huynh-Feldt</th>
<th>Lower-bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>0.325</td>
<td>73.768</td>
<td>5</td>
<td>.000</td>
<td>0.607</td>
<td>0.623</td>
<td>0.333</td>
</tr>
</tbody>
</table>
Table 12

*Test of Within-Subjects Effects*

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>24.154</td>
<td>3</td>
<td>8.051</td>
<td>.252</td>
<td>.860</td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>24.154</td>
<td>1.821</td>
<td>13.262</td>
<td>.252</td>
<td>.757</td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>24.154</td>
<td>1.869</td>
<td>12.921</td>
<td>.252</td>
<td>.763</td>
</tr>
<tr>
<td>Lower-bound</td>
<td>24.154</td>
<td>1.000</td>
<td>24.154</td>
<td>.252</td>
<td>.618</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>6432.390</td>
<td>201</td>
<td>32.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>6432.390</td>
<td>122.021</td>
<td>52.716</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>6432.390</td>
<td>125.246</td>
<td>51.358</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-bound</td>
<td>6432.390</td>
<td>67.000</td>
<td>96.006</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 13

*Test of Between-Subjects Effects*

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>684681.028</td>
<td>1</td>
<td>684681.028</td>
<td>2121.148</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>21626.791</td>
<td>67</td>
<td>322.788</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sphericity is an alternative measure of homogeneity of variance in between-group ANOVA. Sphericity assesses whether the effects of an experimental conditions are consistent among the subjects. The sphericity hypothesis states that the variance of
the difference scores in a within-subjects design are equal across groups. Mauchly's test is an assessment of the sphericity assumption using the chi-square test (Leech et al., 2005). Unfortunately, Mauchly's test is somewhat problematic (Leech et al., 2005). For small sample sizes the test tends to overlook sphericity violations resulting in Type II errors (Sproull, 2002). When sample sizes are large, the test tends to indicate erroneous significant results resulting in Type I errors (Kesselman, Rogan, Mendoza, and Breen, 1980; Sproull, 2002). Therefore, results of Mauchly’s test may be misleading when determining whether the sphericity assumption has been violated. For this pilot study, the hypothesis of sphericity was not rejected for within subjects effects ($p > .05$), Sig. = .000. The sphericity assumption was met.

The test of within subjects effects determines whether the means are equal. From the repeated-measure ANOVA test of within subjects effects table, $F-calculated = .252$, df(3, 201), $\alpha = .05$ and $p < .001$. From the $F$-distribution table, $F-critical = 2.6049$. $F-critical$ was $> F-calculated$ and Sig. values were greater than .001. The null hypothesis that within-subject means were not significantly different cannot be rejected. These results were not statistically significant.

From the test of between subjects effects table, $F-calculated = 2121.148$, df(1, 67), $\alpha = .05$ and $p < .001$. From the $F$-distribution table, $F-critical = 4.0012$. $F-critical$ was $< F-calculated$ and Sig. $< .001$. The null hypothesis that between-subject means are equal was rejected. These results were statistically significant.

To test the hypothesis of stability and replicability of the study, confidence intervals were computed for the 6 predictor variables at the 95% level. These intervals
were computed from raw scores and are indicated in Figure 2. In the case of each variable, the confidence interval captures the mean.

Hierarchical cluster analysis, also referred to as segmentation analysis or taxonomy analysis, has been defined as a dimension-free classification method for partitioning a multivariate, heterogeneous sample into relatively homogeneous sub-samples or clusters (Afifi, Clark and May, 2004; Berenson, Levine and Goldstein, 1983; Diekhoff, 1992; Everitt, Landau and Leese, 2001). The goal of cluster analysis is to draw out nested clusters from a sample such that members of a particular cluster are more similar to entities within that cluster than entities in other clusters (Anderberg, 1973). The outcome of cluster analysis is a classification scheme by which objects are divided into subgroups (Anderberg, 1973). While there is no standard definition of the concept of clusters, clustering is important in all branches of science because all domains seek to characterize and classify their objects of study (Der and Everitt, 2001). Among the principle applications of cluster analysis are hypothesis generation and hypothesis testing (Aldenderfer and Blashfield, 1984). The Pearson and Spearman coefficients are often used in cluster analysis to measure similarity and dissimilarity among cluster members (Kaufman and Rousseeuw, 1990).

A hierarchical agglomerative cluster analysis was performed using SPSS version 10 to ascertain whether significant differences were present between the treatment and the control (Afifi et al., 2004; Aldenderfer and Blashfield, 1984; Anderberg, 1973; Der and Everitt, 2001; Diekhoff, 1992; Kaufman and Rousseeuw, 1990). The goal of the analysis was to test the hypothesis that skills transfer and knowledge growth occurred between tests as demonstrated by changes in test scores (Afifi et al., 2004; Aldenderfer
<table>
<thead>
<tr>
<th>Variable</th>
<th>CI Upper Bound</th>
<th>CI Lower Bound</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>testa1</td>
<td>75.720</td>
<td>69.950</td>
<td>72.840</td>
</tr>
<tr>
<td>testa2</td>
<td>151.440</td>
<td>139.910</td>
<td>145.680</td>
</tr>
<tr>
<td>testb1</td>
<td>71.140</td>
<td>63.540</td>
<td>67.340</td>
</tr>
<tr>
<td>testb2</td>
<td>158.340</td>
<td>138.550</td>
<td>148.440</td>
</tr>
<tr>
<td>quiz</td>
<td>128.030</td>
<td>121.090</td>
<td>124.560</td>
</tr>
<tr>
<td>lab</td>
<td>204.220</td>
<td>191.870</td>
<td>198.040</td>
</tr>
</tbody>
</table>

*Figure 2. Confidence intervals for 6 predictors variables.*
Figure 3. Dendrogram generated with SPSS version 10.
and Blashfield, 1984; Anderberg, 1973). Validity of the analysis is demonstrated in that
differences between test scores are apparent. No standardization of the data matrix was
performed because data were already in standard score format. The resemblance
coefficient used was Euclidean distance because it has an upper bound, is scale
dependent and is the most often used coefficient employed for this pilot study was
nearest neighbor using Pearson correlations as intervals (Aldenderfer and Blashfield,
1984; Diekhoff, 1992). A cluster membership range of 3 to 6 was selected and both a
Proximity Matrix and Agglomeration Matrix were generated (Afifi et al., 2004; Dekhoff,
1992.). Entities selected for the cluster analysis were test scores, lab scores and quiz
scores. These variables were selected because they were deemed most representative
of participant performance with and without the treatment. Given the moderate sample
size and number of variables under examination, Q-analysis and R-analysis were
deemed unlikely to yield meaningful results.

Referring to the dendrogram or tree diagram in Figure 3, the horizontal scale
along the top of the diagram indicates the range of Euclidean distance coefficients \( e_{jk} \)
for the clustering procedure (Anderberg, 1973; Romesburg, 1984). For example, the
Euclidean distance coefficient for \( z_{testb1} \) and \( z_{testb2} \) is \( e_{34} = 8.0 \). Each vertical line in
the dendrogram represents a single step in the clustering process where two clusters
are joined (Der and Everitt, 2001; Romesburg, 1984). The length of the vertical line is
positively correlated with the difference between the two clusters being joined (Afifi, et
al., 2002; Berenson et al., 1983). From the dendrogram it can be observed that \( z_{testa1} \)
and \( z_{testa2} \) (clusters 1 and 2) are the most similar among the six variables in the
analysis followed by \( z_{testb2} \) and \( z_{quiz} \) (clusters 4 and 6). Also, \( z_{testb1} \) and \( z_{testb2} \)
(clusters 3 and 4) are similar but significantly less similar than ztesta1 and ztesta2. In addition, the cluster containing zquiz, ztestb1 and ztestb2 is significantly dissimilar from the cluster containing ztesta1 and ztesta2. The single cluster containing the zlab (cluster 5) is the most dissimilar of all clusters formed. The following conclusions were drawn from this analysis.

First, scores on the treatment scales, ztesta1 (pretest) and ztesta2 (posttest), were very similar. Likewise, scores on the control scales, ztestb1 (pretest) and ztestb2 (posttest), were also similar but significantly less so than on the treatment scale. These data indicate very little difference between scores on ztesta1 and ztesta2 but significant difference between scores on ztestb1 and ztestb2. Something changed between the treatment and the control as well as between ztestb1 and ztestb2.

Second, the significant dissimilarity between the two clusters containing the treatment and control tests may indicate that transfer of knowledge and increased skills mastery occurred between the two sets of tests. The same can be said for the dissimilarity between scores on ztestb1 and ztestb2. These data support the hypothesis that a difference in test scores from the treatment to the control can be attributed to at least some degree to the presence of the simulation as the treatment.

Third, the zlab score was significantly different from the other 5 variables measured. Based on the data, we may conclude that the zlab score was effected by all 4 test scores as well as the zquiz score. While the presence of the simulation may have contributed to this dissimilarity, it is not possible to determine the degree of this contribution. In this instance, clear support for the research hypothesis is not evident.
Cronbach’s alpha is defined as a coefficient of reliability or consistency. For this pilot study, a Cronbach’s alpha was computed to assess whether the six variables, testa1, testa2, testb1, testb2, lab and quiz formed a reliable measurement scale. The alpha for these items was 0.8735, which indicated that the items formed a scale that was reasonable exhibiting internal consistency and reliability. Five of the six variables exhibited high corrected item total correlations. Lab exhibited a low correlation of 0.2665 indicating it did not constitute a good component. Removing lab from the scale yielded a higher alpha of 0.9157.

Conclusions

A one-way ANOVA was computed to test the homogeneity of variance assumption. A Sig. of 0.253 > .05 from Levene’s test was not significant indicating that variances among the six groups of scores (testa1, testa2, testb1, testb2, lab and quiz) was not significantly different. The assumption of homogeneity of variance was met. This conclusion was further supported by results from Mauchly’s test of sphericity. In addition, the confidence intervals for all 6 predictors captured their respective mean values indicted in Figure 2.

Results of a paired-samples t-test did not support the research hypotheses. This test did not reveal any significant difference in test scores between the treatment and control phases of the pilot study.

Pearson r values for the 4 tests and the lab grade where weak; testa1 and lab \((r(65) = 0.336, p < .001)\), testa2 and lab \((r(65) = 0.358, p < .001)\), testb1 and lab \((r(65) = 0.143, p < .001)\) and testb2 and lab \((r(65) = 0.158, p < .001)\). Interaction with the simulation did not appear to positively impact student lab scores.
Results obtained from the MLR were statistically significant indicating that the variables testb1 and testb2 were good predictors of course grade. This conclusion was supported by the very high positive relationship between testb2 and course grade, ($r(65) = 0.915$, $p < .001$). In addition, the effect size of the combined predictor variables was strong. However, the fact that testb1 and testb2 were better predictors of course grade than testa1 and testa2 further indicated that the research hypothesis was not supported.

Students were presented with the same questions on both testa1 and testa2 but with the order of questions changed. The same procedure occurred with testb1 and testb2. This procedure was followed to ensure good test-retest reliability and criterion-related validity.

Student performance on testa1 and testa2 were very similar indicating that these scores were highly correlated. Performance on testb1 and testb2 were also very similar once again indicating a high degree of correlation. Student performance on testb1 and testb2 were also highly correlated with course grade. Quiz and crsgrad also appeared highly correlated. Data from the cluster analysis indicated very little difference between scores on testa1 and testa2 but significant difference between scores on testb1 and testb2. The significant dissimilarity between the two clusters containing the treatment and control tests may indicate that transfer of knowledge and increased skills mastery occurred between the two sets of tests. The same can be said for the dissimilarity between scores on testb1 and testb2. These data support the hypothesis that a difference in test scores from the treatment to the control can be attributed to at least some degree to the presence of the simulation as the treatment.
Of 68 participants in the pilot study, only 19 participants (27.94%) completed the self-reporting knowledge-seeking survey. This low response rate rendered data collected from the survey anecdotal at best and of minimal analytical value. Low response was attributed to the time interval during which students were asked to complete the survey. During the final week of the fall 2006 term students were concentrating on final exams and completing semester projects. Many students may have viewed the survey as just one more thing consuming valuable study time. Nevertheless, some results from the survey are worthy of note. First, more than 50% (52.6) of respondents reported having previously taken at least one science course at the university level. These students may have experience a high degree of comfort with the course content and the simulation than their counterparts possessed of no science background at the university level. Second, slightly more than 15% of respondents reported engaging in at least one form of knowledge seeking behavior outside of class. While it is good to see that some students are willing to do extra work outside of class, this response was too small to allow extrapolation of cause and effect.

Support for the research hypothesis was not significant based on result from the pilot study. The fact that testb1 and testb2 were better predictors of course grade than testa1 and testa2 may indicate that students were able to transfer skills acquired through exposure to the simulation to a similar but non-equivalent domain assessed by testb1 and testb2. Cluster analysis results seem to support this conclusion. However, students did not perform significantly better on testa2 than on testa1 indicating that interaction with the simulation had little immediate impact.
REFERENCES


Lazarowitz, R. and Huppert, J. (Spring 1993). Science process skills of 10th grade


Shadish, W.M., Cook, T.D. & Campbell, D.T. (2002). Experimental and quasi-


