ITERATIVE USABILITY TESTING IN THE DEVELOPMENT OF A LEARNING TECHNOLOGY SYSTEM FOR TEACHING GEOGRAPHIC INFORMATION SYSTEMS WITHIN A CIVIL ENGINEERING CURRICULUM

by

JOHN MCKENNA SULLIVAN

A THESIS

Presented to the Faculty of the Graduate School of the UNIVERSITY OF MISSOURI-ROLLA

In Partial Fulfillment of the Requirements for the Degree MASTER OF SCIENCE IN INFORMATION SCIENCE & TECHNOLOGY

2004

Approved by

_______________________________           _______________________________
Richard H. Hall, Advisor                     Madhu Reddy

_______________________________
Ronaldo Luna
ABSTRACT

The goals of this project were to: a) carry out the evaluation of a learning system for teaching Civil Engineering students to use Geographic Information Systems software within the context of relevant problems and b) explore the efficacy of iterative usability testing as a tool for development of learning technologies. A series of three usability tests were conducted based on an initial proposed design of the learning system. The first usability test was conducted on a pre-existing system with a design similar to that proposed, while the other two tests were conducted on an initial and production prototype respectively. In general, the results pointed to the role of the learning system as a support tool, and a consequent design that focuses on progressive scaffolding, flexibility, and logical consistency. The results also supported the efficacy of usability testing, in that iterations of the system improved as development progressed, informed by the iterative usability testing. Finally, the importance of scarce resources and education of the development team on the role of usability testing, were identified as central factors in determining the effectiveness of usability testing.
ACKNOWLEDGMENTS

I would like to express my deepest appreciation to those who have not only guided my efforts in this endeavor but also in the path to self-discovery. My advisor Dr. Hall- a man I will be indebted to for long time to come. His continuing advice and support has been constant. His subtle guidance delivered with infinite patience shall not be soon forgotten.

I would like to than my co-advisor, Dr. Madhu Reddy. His “tough love” provided me stalwart convictions and made me contemplate otherwise forgotten questions. Many eye-opening discussions and good times were afforded in his classroom and office. He has shaped my education since he arrived to the Academia that is University of Missouri – Rolla, and I know he will continue to do so for others.

Dr. Ronaldo Luna deserves my utmost gratitude expressed- his never ceasing drive towards perfection is tangible and contagious. His support has been invaluable to the projects development. His spirit motivated numerous times during the course of the project; while my mind would be suffering from fatigue, his would merely be starting.

Special thanks go to my parents, family, and friends; all of whom has had to put up with me over the past few months. Without their support and encouragement none of this would have been possible.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1. OVERVIEW</td>
<td>1</td>
</tr>
<tr>
<td>1.2. FORMATIVE-ITERATIVE EVALUATION METHODS WITH USABILITY TESTING</td>
<td>2</td>
</tr>
<tr>
<td>1.2.1. Overview</td>
<td>2</td>
</tr>
<tr>
<td>1.2.2. Prototyping</td>
<td>3</td>
</tr>
<tr>
<td>1.2.3. Usability testing</td>
<td>6</td>
</tr>
<tr>
<td>1.3. FORMATIVE-ITERATIVE EVALUATION METHODS IN EDUCATION</td>
<td>7</td>
</tr>
<tr>
<td>1.4. PROJECT CONTEXT</td>
<td>10</td>
</tr>
<tr>
<td>1.4.1. Application: Geographic Information Systems for civil engineering</td>
<td>10</td>
</tr>
<tr>
<td>1.4.2. Learning objects and SCORM compliance</td>
<td>12</td>
</tr>
<tr>
<td>1.4.3. Progressive scaffolding</td>
<td>13</td>
</tr>
<tr>
<td>1.4.3.1 Previous Experiment 1: pilot exploration of different types media scaffolds</td>
<td>13</td>
</tr>
<tr>
<td>1.4.3.2 Previous Experiment 2: comparison of video vs. static graphics as scaffolds</td>
<td>14</td>
</tr>
<tr>
<td>1.5. RESEARCH GOALS</td>
<td>16</td>
</tr>
<tr>
<td>2. METHOD AND RESULTS</td>
<td>17</td>
</tr>
</tbody>
</table>
2.1. IDENTIFICATION AND TESTING OF AN EXISTING SYSTEM .......... 17
  2.1.1. Overview ........................................................................................................ 17
  2.1.2. MELL Testing .................................................................................................. 20
    2.1.2.1 Participants ................................................................................................. 20
    2.1.2.2 Directions .................................................................................................... 20
    2.1.2.3 Tasks ............................................................................................................ 21
    2.1.2.4 Post-testing questionnaire: ........................................................................ 22
  2.1.3. Results ........................................................................................................... 22
    2.1.3.1 Quantitative Results .................................................................................... 22
    2.1.3.2 Open Ended Questionnaire Responses ..................................................... 23
    2.1.3.3 Summary of observations from field notes .............................................. 24
  2.1.4. Discussion ...................................................................................................... 26
  2.1.5. Recommendations .......................................................................................... 27

2.2. INITIAL PROTOTYPE USABILITY TESTING ............................................ 27
  2.2.1. Overview ........................................................................................................ 27
  2.2.2. Participants & pre-questionnaire ................................................................. 29
  2.2.3. Experiment design ......................................................................................... 29
  2.2.4. Tasks ............................................................................................................... 30
  2.2.5. Post-testing questionnaire ............................................................................. 31
  2.2.6. Results ............................................................................................................ 31
    2.2.6.1 Questionnaire and task completion tables ................................................ 31
    2.2.6.2 Open ended questionnaire responses ....................................................... 32
    2.2.6.3 Overall observations as recorded in field notes ....................................... 33
2.2.7. Discussion .............................................................................................. 35
2.2.8. Recommendations .................................................................................. 37

2.3. FINAL PROTOTYPE USABILITY TESTING............................................... 38

2.3.1. Participants & pre-questionnaire ............................................................ 38
2.3.2. Experiment design .................................................................................. 38
2.3.3. Tasks ....................................................................................................... 39
2.3.4. Post-testing questionnaire ....................................................................... 42
2.3.5. Quantitative Results ............................................................................... 43
2.3.6. Qualitative Results ................................................................................. 44

2.3.6.1 Subject Recommendations from Exit Questionnaire ......................... 44
2.3.6.2 Overall observations as recorded in field notes .................................. 45
2.3.7. Discussion .............................................................................................. 48
2.3.8. Recommendations: ................................................................................. 49

3. DISCUSSION & CONCLUSION ............................................................................ 51

3.1. OVERVIEW ..................................................................................................... 51

3.2. OBJECTIVE 1: EVALUATING MODULE FOR PEDAGOGICAL EFFECTIVENESS ........................................................................................................ 51

3.2.1. Overview ................................................................................................ 51
3.2.2. Usability testing of an existing system ................................................... 51
3.2.3. Usability testing of initial prototype ....................................................... 52
3.2.4. Usability testing of production prototype ............................................... 53
3.2.5. General Insights and Design Recommendations .................................... 55

3.3. EXPLORE THE EFFICACY OF ITERATIVE USABILITY TESTING AS A TOOL FOR DEVELOPING LEARNING TECHNOLOGIES .................. 56
3.4. FUTURE WORK............................................................................................................. 57
3.5. CONCLUSION.................................................................................................................. 58

APPENDIX A: Subject Explanations for Questionnaire (Entrance & Exit) Responses for MELL Existing System.......................................................................................................................... 62

APPENDIX B: Subject Explanations for Questionnaire (Entrance & Exit) Responses for Initial Prototype ........................................................................................................................................ 66

APPENDIX C: Subject Explanations for Questionnaire (Entrance & Exit) Responses for Final Prototype.................................................................................................................................... 70

APPENDIX D: Technically Oriented Questions Administered Concurrently with the Final Prototype Exit Questionnaire.......................................................................................................................... 74

BIBLIOGRAPHY .......................................................................................................................... 77

VITA ........................................................................................................................................ 80
# List of Illustrations

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1 - User Centered Design</td>
<td>3</td>
</tr>
<tr>
<td>Figure 1.2 - Learning System Model</td>
<td>11</td>
</tr>
<tr>
<td>Figure 2.1 - MELL Website Text and Video Demonstration</td>
<td>18</td>
</tr>
<tr>
<td>Figure 2.2 - ArcView Learner System after Initial Prototype Recommendations Implemented</td>
<td>28</td>
</tr>
<tr>
<td>Figure 2.3 - Screenshot of video demonstrating ArcView fundamental after Initial Prototype Recommendations Implemented</td>
<td>28</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1 - Relative Effectiveness of Low vs. High-fidelity Prototypes (Rudd 1996)</td>
<td>4</td>
</tr>
<tr>
<td>Table 2.1 - Quantitative Testing Results of MELL</td>
<td>23</td>
</tr>
<tr>
<td>Table 2.2 - Quantitative Testing Results for Initial Prototype</td>
<td>32</td>
</tr>
<tr>
<td>Table 2.3 - Final Prototype Testing Results</td>
<td>44</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

1.1. OVERVIEW

This research was part of a proof-of-concept project funded by the National Science Foundation’s CCLI (NSF 03-558) Educational Materials program. The ultimate goal of this project was to develop a web-based learning system that will teach students the use of Geographic Information Systems (GIS) within the foundational courses of a typical Civil Engineering program. As opposed to generating a series of GIS courses, the GIS know-how was introduced within an existing course as a module that will reinforce basic concepts taught in a comprehensive manner. The prototype being developed in this proof-of-concept project has a geotechnical emphasis for inclusion in a typical soil mechanics course. Based on the results from the proof-of-concept project, the prototype will serve as a template to further develop other modules (environmental, transportation, hydrology, and/or construction) in a potential full development effort involving other faculty in the Civil, Architectural and Environmental Engineering Department. The principal objectives of this project are:

a) To create a prototype web-based learning system to support students in learning how to apply GIS within the context of Civil Engineering (geotechnical emphasis) course CE 215 Fundamentals of Geotechnical Engineering
b) To carry out an initial series of evaluation studies with components and iterations of this system.

c) To use the prototype system and evaluation data as the foundation for full scale, comprehensive project proposals.

Currently, the first two objectives have been accomplished. A web-based learning system was created with the civil engineering context and initial series of evaluations were completed on the system. This evaluation research is the focus of this thesis. System designs were evaluated by usability testing in three separate iterations: pre-design- where an existing system with similar characteristics was identified and tested, initial design- where the early design choices were partially implemented and tested, final design- where the final design produced was tested and tweaked. The final design testing was intended, in part, to ensure that the final prototype was ready for full scale testing.
This work focuses on the preliminary iterations of usability testing and how they helped improve the design and effectiveness of the prototype. Before these two points are discussed further, a background of usability testing, usability testing in education, SCORM, progressive scaffolding, and a deeper level of project detail will be discussed in the following sections.

1.2. FORMATIVE-ITERATIVE EVALUATION METHODS WITH USABILITY TESTING

1.2.1. Overview  Formative evaluation, the process of evaluating designs early during product development, was not always considered important to design as it is today. When usability was first introduced as a method for improving design, many companies would simply (summative) “test” their product just prior to shipping- a different type of quality assurance (Barnum 2002). Testing was done to determine that their product was good and provide them the chance to claim that it is user-tested, human-factor satisfied, or ergonomically sound. The problem that many companies faced was that their testing often did not reveal these results. Users would struggle with features and their implementations, and a particular design would be discovered to have a plethora of flaws. However, by this stage of development redesign would be to costly, and imperfect designs were delivered with improvements being stored for future implementations. As a result of these previous outcomes, more companies are embracing formative evaluation coupled with summative evaluation as a gateway towards user-centered design. Products and their prototypes are being tested with users much earlier in the design cycle, often before any physical design is developed. The Standish Group analyzed factors predicting the successfulness of software development and discovered that user involvement is the number one predictor (Kreitzberg 1999). By employing formative evaluation methods early and repeatedly with users, one can be more assured that the criterion of user involvement is satisfied to a higher degree. Formative evaluation involves users and allows the design team to learn from them. These evaluations are categorized into one of three phases, early, middle, or late, where the distinction is derived from the fidelity of the model or prototype to be evaluated.
In the modern usability analysis, usability testing and prototyping are interconnected. Prototypes must be developed in order to experiment with different designs and features, explore the feasibility of different aspects, but also to allow interaction with end users with some form of analysis. Usability testing offers an avenue to perform analysis on a given system’s prototype. Prototyping combined with usability testing creates an adaptive design cycle and encourages formative iterative evaluation methods. The design errors learned from the first batch of testing leads to a further developed prototype. This new prototype of design is now ready for testing again and can begin the cycle anew (see Figure 1.1). To further understand the interrelation between prototyping and usability testing, both subjects will be examined in more detail in the following sections.

1.2.2. **Prototyping** Prototyping is a powerful tool that can be implemented in formative evaluations to discover benefits and problems of a particular design when combined with iterative usability testing. Prototyping is the process of designing a mock-up of some
specified degree for one or more of the following purposes: to test the feasibility of an idea, clarify requirements, or allow user testing/evaluation (Preece 2002). In essence, prototypes provide the content to be tested for traditional usability testing. A user will interact with a prototype representative of the current level of development and through the use of usability testing, necessary design changes can be made for the next iteration. Generally, two levels of prototypes are used as descriptors: low-fidelity and high-fidelity prototypes. Low-fidelity prototypes are generally low in cost, time, and/or complexity to create, and are therefore used in the early (formative) phases of design with the more complex high-fidelity prototypes used in more summative testing. A more exhaustive list of advantages and disadvantages is provided below:

Table 1.1 - Relative Effectiveness of Low vs. High-fidelity Prototypes (Rudd 1996)

<table>
<thead>
<tr>
<th>Type</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-Fidelity</td>
<td>Lower Development costs</td>
<td>Limited error checking</td>
</tr>
<tr>
<td></td>
<td>Evaluate multiple design concepts</td>
<td>Poor detailed specifications to code to</td>
</tr>
<tr>
<td></td>
<td>Useful communication device</td>
<td>Facilitator driven</td>
</tr>
<tr>
<td></td>
<td>Address screen layout issues</td>
<td>Limited utility after requirements established</td>
</tr>
<tr>
<td></td>
<td>Useful for identifying market requirements</td>
<td>Navigation and flow limitations</td>
</tr>
<tr>
<td></td>
<td>Proof-of-concept</td>
<td></td>
</tr>
<tr>
<td>High-Fidelity</td>
<td>Complete functionality</td>
<td>More expensive to develop</td>
</tr>
<tr>
<td></td>
<td>Fully interactive</td>
<td>Time-consuming to create</td>
</tr>
<tr>
<td></td>
<td>User-driven</td>
<td>Inefficient for proof of concept designs</td>
</tr>
<tr>
<td></td>
<td>Clearly defines navigational scheme</td>
<td>Not effective for requirements gathering</td>
</tr>
<tr>
<td></td>
<td>Use for exploration test</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Look and feel of final product</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Serves as a living specification</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marketing and Sales tool</td>
<td></td>
</tr>
</tbody>
</table>
Given these attributes, it is not hard to see why low fidelity prototypes work well with formative phases of design, where money and time for testing may be hard to justify. Design errors discovered from these prototypes tend to be large in scale, and can save countless hours and dollars since the design has yet to be truly developed. Examples of such prototypes include: paper sketches, storyboarding, HTML, and other low production mock-ups. In contrast, high-fidelity prototypes are more costly, require more time, and have greater complexity than their low-fidelity counterparts. The advantage high-fidelity prototypes bring to design is that they are approaching the final design, or are a subsection thereof. Details of a particular design can be tested more accurately, and fine tuning of a system can occur. The drawbacks of relying solely on high-fidelity prototypes are quite dubious though; they take long time to build, testers tend to comment on superficial aspects versus content, developers are disinclined to make alterations after investing their time, software prototypes can set the expectations too high, and a single bug in the prototype can force testing to stop (Dumas and Redish 1999). Some of these drawbacks of high-fidelity prototypes can be alleviated by using a fully functioning system that already exists but attacks many of the design concerns that one’s project may be facing. By performing testing on such a prototype, one can reap the benefits of the high-end prototype testing, but without having to invest the time required to craft one. Even without such a system, this is not to say that summative testing with high-fidelity prototypes is not worthwhile; rather that one needs to be careful when doing so and makes sure the prototype fits the needs of the next iteration of testing and design. A prototype that does not attempt to answer the design questions of the current phase of development provides little assistance. Prototypes are required for usability testing, since the user needs to be able to test something, no matter how detailed.

Even after deciding on the required fidelity of a prototype, one needs to consider two other factors, features and functionality, and how they relate to the evaluation before building a prototype. When designing prototypes, time and cost restrictions force one, if design questions alone do not, to make trade-offs between the number of features (breadth) and functionality (depth), resulting in a partial prototype. Three general classifications of such prototypes are: Horizontal, Vertical, and Scenario (Nielsen 1989). Horizontal prototypes contain many functions of the design/interface, but to a very
limited depth. An example of a software-based horizontal prototype would be where all the menu buttons are developed to the first level, so a user could click on it and go to a sub-menu but no further. The vertical prototype is the opposite of the horizontal, because it selects a few features and develops them to a much deeper extent, perhaps to completion. An example of this type of prototype could be a software prototype with only one main menu button (File perhaps), but with all of its sub-menus and functions close to or fully operational. The third classification, scenario, is a task-oriented combination of horizontal and vertical prototypes. Instead of developing a prototype with the purpose of a great amount of features or functionality, this prototype is developed to finish a certain, generally low-level, task. All features and functionality required to complete a given task are implemented. This characteristic makes them an obvious choice for evaluating usability, since users are given real tasks instead of being told to complete a set of unrelated transactions. Given the limited size of these prototypes, implementing changes based on feedback is relatively easy. While partial prototypes do offer a plethora of benefits, there are some difficulties as well. A prototype which is not fully implemented, where all user paths are developed, limits the number of ways a user can actually fail (Dumas and Redish 1999). A user might try an undeveloped avenue of the prototype and quickly discover that it was in fact the wrong direction for the given testing. When the final design is produced, however, all avenues of the product are available, and a user who went the wrong way towards completing his/her goal could travel increasing farther away from it. Care must be taken to not over assume the usability of a design from the results of usability testing in such a situation.

1.2.3. **Usability testing** Usability testing is the process of observing users interacting with and using a product to ascertain how well users function with the product. The iterative nature of usability testing comes when the testing is integrated into the design and development cycle resulting in incremental improvements. Characteristics that are generally found in usability testing (Dumas and Redish 1999) are:

1. Each test has particular goals and concerns stated when designing the test
2. The participants are representative of real users
3. Participants work on real tasks
4. Participants actions and words are recorded and observed
5. Data is analyzed to diagnose problems and propose recommendations

Thus, in usability testing there is a controlled situation, recognizable and explicit goals, and nearly everything (times, comments, pauses, expressions, etc) is recorded. Usability testing is generally considered to be different than a traditional experiment; the latter consisting of formal hypothesis testing, the manipulation of an independent variable, and the random assignment of participants to experimental groups while usability testing is closer in nature to observational or exploratory research, where variables are not manipulated, and hypotheses are not required. Questionnaires and interviews are often given to determine prior knowledge as well as satisfaction with the product. Users are often asked to adhere to a “talk-aloud” protocol, the process of explicitly stating what one is thinking during a given task, to provide rich data about ongoing thought processes. Usability testing typically employs a small number of users, a phenomenon referred to as “discount usability testing,” because this results in a discount in both cost and time (Nielsen 1994). Due to the low number of participants, there is generally an emphasis on qualitative, rather than quantitative analyses, and inferential statistics with an emphasis on probability levels are not usually applied. Although this precludes the ability to perform inferential statistics, and assess statistical significance, qualitative analysis can offer important insights based on the plethora of data available for each participant. These in-depth qualitative analyses with a small number of users are particularly well suited for formative evaluation, which refers to evaluation occurring during the early stages of the design and development cycle. These low-cost usability tests can be applied iteratively with system prototypes. Even though the methods of prototyping and usability testing have proved themselves, their implementation in educational learning technologies has been limited.

1.3. FORMATIVE-ITERATIVE EVALUATION METHODS IN EDUCATION

Some researchers who work within both the Human-Computer Interaction (HCI) and Learning Technology communities have suggested that the communication between the two is less than optimal, despite the fact that they share common subject matter (Dillon 2000). Though usability testing of web-based learning technologies has become
more common in recent years (Dean 1999; Richardson, Swan et al. 2000), it is still relatively rare. There are a number of reasons why usability testing has not become a standard tool for learning technology researchers. First, usability testing does not fit well within the most common conceptions of qualitative research in education, which tend to focus more on naturalistic observation, with intense prolonged contact in a field setting (Miles and Huberman 1994). While usability metrics can be gauged with ethnographic methods such as field studies, the time and resources needed to devote towards such endeavors are substantial when compared to standard usability testing, as applied in industry. Second, the small number of participants typically used in usability studies does not allow for much generalization of results, particularly if the technology being studied has a wide and diverse effect on users. The small number of participants required for usability studies precludes quantitative/statistical analysis of the data, since the number of users does not provide a statistical significance. Third and conversely, the use of single-subject observation and intensive data collection and analysis does not easily allow for the large number of participants required for any sort of inferential statistics to be applied. Observing users and the responses they give in questionnaires as well as the “think-aloud protocol” provides a plethora of data even for limited number of users. This mountain of information grows seemingly exponentially in relation to the number of users, and the amount of qualitative data created by the number of subjects required for proper quantitative analysis is simply unwieldy. Fourth, and most fundamentally, usability testing is usually implemented with well-defined tasks and well-defined performance measures. Usability tasks typically involve activities such as searching for a given piece of information, entering data, or manipulating a program in a specific manner. Educational researchers, on the other hand, are often interested in outcomes such as knowledge integration, problem solving, and critical thinking.

Despite these differences in approach, they are not mutually exclusive. Both inquiries are intended to produce a better product; and, to some extent, they simply address different levels of the same problem. Educational research tries to determine whether, or to what degree, a product affords learning opportunities. Their prototypes strive towards making education opportunities abundant and powerful, and are measured on these dimensions. On the other hand, usability assessments try and discern whether,
or to what degree, an interface facilitates an effective interaction. Usability studies/experiments attempt to discern the level any given system or product has in allowing its users to be able to complete their intended goals. In essence, educational testing could determine the pedagogical merit while usability testing could determine how easy that said merit is to achieve. A product might have stellar potential for education, but if it has a complex and confusing interface, that educational potential could be lost. Inversely, a product could be a simple, intuitive tool, but it may not offer any true educational benefits. Thus, the approaches can be synthesized to address different levels of a composite problem. Such a synthesis is discussed in further detail in the following section.

There are also some specific cases where traditional usability testing can be very beneficial for the evaluation of educational technologies. In particular, when used as a supplement to more traditional quantitative experimentation, a detailed analysis of a small number of typical users interacting with instructional software can provide invaluable insights and explanations with respect to the quantitative results (Hall, Philpot et al. 2004). Such a multi-method approach is traditionally referred to as triangulation — a term that is popular within both the HCI and educational research communities. What might be indicated by qualitative measures and tests could be further verified or qualified by either correlating or conflicting evidence discovered by more traditional quantitative investigations. Educational researchers point out that triangulation can be used to check the validity and reliability of findings (O'Malley and Valdez 1996), and that multiple forms of evidence will provide a more accurate picture of the student (Wiggins 1998). By combining qualitative and quantitative studies, one can be more assured that the conclusions made are accurate and truly reflect the reality of the situation. Similarly, human-computer interaction (HCI) researchers Wendy Mackay and Anne-Laure Fayard point out that research effectiveness can be enhanced by triangulation across research methodologies and across academic disciplines (Mackay and Fayard 1997). They reiterate the point that more data sources produce results that are more reliable and able to be generalized. Mackay and Fayard do not stop there however, by discussing the use of methodologies across disciplines, they add another dimension to the triangulation aspect. Triangulating does not just include the types and styles of research occurring, it
also includes the “lens” of interpretation by which the research is scrutinized. Although this paper focuses on formative usability testing as a part of initial development, the system will be assessed via more traditional experimental methods once the prototype system is completed, so the research program will included multiple methodologies and conclusions will be based on the triangulation of multiple sources. This triangulation of data should assist in the quality of results from being compromised.

1.4. PROJECT CONTEXT

1.4.1. Application: Geographic Information Systems for civil engineering  The learning system serves as a focus of the project is rooted in teaching civil engineering students how to use GIS. A Geographic Information System (GIS) is a computerized database management system that provides geographic access (capture, storage, retrieval, analysis and display) to spatial data. GIS can produce maps and pertinent information in a visual format regarding a specific area based upon the spatial data provided. While the industry sector of civil engineering has begun the process of integrating GIS itself, the academic world has been slower to respond (Keshawarz 2001). Since civil engineering is replete with uses for GIS functions, public agencies’ (the civil engineer’s primary employer) use of GIS technology is increasing rapidly. There exists a consequent need for civil engineers versed in GIS and able to apply GIS tools to civil engineering problems in innovative ways.

The learning system developed for the civil engineering curriculum for this project focuses on a geotechnical application. The prototype consists of a comprehensive problem and an associated repository of learning objects organized using progressive scaffolding approach (Sullivan, Hall et al. 2004). Figure 1.2 represents a schematic of the basic system framework. The system consists of three parts, foundational knowledge in civil engineering; foundational knowledge in Arcview®, which is a popular GIS software application; and an applied problem. The problem is the driving force within the system and integrates the civil engineering with the GIS concepts that are intended to be taught. The system was used in University of Missouri – Rolla Civil Engineering 215 class, Soil Mechanics, where students learn civil engineering concepts and have a first order working knowledge of these concepts. The early application of these concepts is important to enhance learning, much like a laboratory experience where concepts are
taught in lecture and applied in the lab. The students’ knowledge of GIS is diverse, because the course where the system is to be used is multidisciplinary with students from various engineering disciplines. Therefore, the GIS modules are an important part of the prototype system. The learning objects are to be organized as scaffolded media; designed so they will be applicable to students with different levels of knowledge.

The applied problem is at the heart of the system, with the GIS learning objects providing support as needed. For the prototype, the basic problem involves simulating the geotechnical engineering decision making in earthwork construction. The basic problem to solve is to select from a variety of spatially distributed soil borrow sites that can meet the construction and engineering objectives at a construction site.

Two important educational and technical criteria guide development: a) The content is broken into re-usable learning objects; and b) The system is compliant with ADL/SCORM guidelines. These criteria are discussed in the next section.
1.4.2. Learning objects and SCORM compliance  The goal of distributed learning networks is to provide a repository of sharable learning objects, facilitated by information networks. Conceptually, this means that educators decompose their courses into a collection of fundamental elements, called learning objects, and make them available to an information network (Committee 2002). A learning object is a collection of web displayable material that has an associated learning objective. There are several goals to such a system. For the objects themselves, it is desired that they be interoperable, accessible, durable, and reusable (Englebrecht 2003).

Key to the success of a distributed learning environment is having a common architecture shared across the network to ensure the interoperability and accessibility of the learning objects. In 1999, Executive Order 13111 tasked the Department of Defense (DoD) “to develop common specifications and standards for technology–based learning” ((ADL) 2004) resulting in the first draft of the Sharable Content Object Reference Model via the DoD’s Advanced Distributed Learning Initiative. The fundamental idea behind SCORM is to use XML to attach metadata tags to content objects. Using these tags, information networks can access and distribute learning objects to a variety of educational environments using Learning Management Systems (LMS).

The primary user of SCORM–compliant distributed learning networks has been the military. The Army has seen remarkable success with its Distributed Learning System (Chisholm 2003), with cost savings resulting in millions of dollars. However, university educational information networks have been slow to adopt and utilize these standards (Cheese 2003). One hindrance is that professors are reluctant to view themselves as “content–providers.” Another fundamental difference between military and academic use is that military applications are more focused on training whereas academic applications are more focused on education. While training tends to emphasize acquiring skills through rote memorization and repetition, education seeks to carry the student through a guided exploration leading toward self-discovery. Since the concept of learning objects has been utilized primarily by the military, most of the key ideas associated with object formation, such as granularity and classification, are more appropriate for training. The GIS project is an ideal translation project because it is a mixture of education and
training. At the training level, the learning objects need to instruct the civil engineering student in the use of the GIS software. At the education level, the learning objects need to provide students with a more complex understanding of the application of GIS as a tool to address a complex problem. Hence, the lessons learned from the military’s development of learning objects can provide a basis for evolving their role into becoming more educationally-oriented.

1.4.3. Progressive scaffolding The term “progressive scaffolding” (Hall 2002) is a term used to refer to a systematic method of providing users with an optimal level of assistance. Within such a system, different levels or tiers of facilitation are provided to match the optimal levels of assistance required. The level could be set by the learner, an instructor, or automatically, based on learner response.

It’s important to note that, scaffolding, as defined within our framework, refers to guidance that supports the core content, which remains constant across differing levels of scaffolding. Therefore, degree of scaffolding is not equivalent to difficulty of the content; rather it refers to the degree of supportive context provided. More specifically, in our research so far, the scaffolding dimension has been represented by the media in which the content is embedded: plain text, text with graphics, or video. Thus the scaffolding differs in the degree of abstraction, fidelity, and richness.

1.4.3.1 Previous Experiment 1: pilot exploration of different types media scaffolds As one part of the learning object repository initiative described earlier, a series of learning modules based on the progressive scaffolding approach as support for a class in Web Development and Design were developed. As a part of this system, a prototype module was created: a step-by-step description for how to create a fairly elaborate web page, using the web development tool Macromedia Dreamweaver®. To create the page, the user must apply a number of general procedures including: setting up a site; adding tables & graphics; using tables for page layout; inserting text; creating hyperlinks; creating image rollovers; and creating a disjoint swap image behavior.

This initial prototype system is straightforward and was intended to examine the key components of the learning system model posed above. The progressive scaffolding is provided in the form of different levels of information for displaying each step in the
development process: a) Text; b) Graphics; c) Narrated Video. The system used for the pilot experiment can be viewed at: http://campus.umr.edu/lite/web_dev_experiment.

A pilot study was carried out using this system including detailed quantitative and qualitative analysis of participants using the system to create a basic web site (Hall, Digennaro et al. 2002; Hall, Digennaro et al. 2003). This pilot study addressed the following two questions:

*To what extent do users utilize the different scaffolding options (text, graphics, and video)?*

*How does their use of the various options relate to performance?*

With respect to our basic experimental questions, it was found that users primarily utilized the most minimal (text) and elaborate (video) scaffolds, while they largely ignored the static graphics. In addition, the amount of time spent viewing the text was positively related to performance; while the time spent viewing the videos was negatively related to performance. Other important findings emerged as well, independent of the experimental questions, based on the qualitative analysis. One of the strongest findings was that the previous content knowledge (experience level) of the user played an important role, with the more experienced users utilizing the more minimal scaffolds to a greater degree. This helped to partially explain the strong negative relationship between the amounts of time spent viewing the videos and performance. The videos most likely did not “cause” poor performance, but were more useful to those with less experience. Unfortunately, this experiment did not allow for adequate evaluation of static graphics as scaffolds, since these were accessed so infrequently. Such an evaluation would have important practical implications, given differences in resources and complexity associated with the development and display of static graphics versus video.

### 1.4.3.2 Previous Experiment 2: comparison of video vs. static graphics as scaffolds

A second experiment was conducted, which extended the first in a number of ways (Hall, Stark et al. 2004). First, a new prototype tutorial was created based on a newer version of Dreamweaver© (http://campus.umr.edu/lite/scaffolding2). In addition, the task participants were to perform was more self-contained, focused, and realistic than the
previous experiment. Second, the number of students who participated was almost tripled (from 7 to 20). Third, and most fundamental, two experimental conditions were created, so that participants either used graphics as an adjunct scaffold to the text, or they used a video, rather than having the option to use both. This allowed for the direct comparison of graphics versus video as a scaffolding technique. Fourth and final, a pre-questionnaire was used which assessed participants’ experience with Dreamweaver© in particular and web development in general, which allowed for a more thorough examination of the role of user experience in it’s impact on the users’ use of the tutorial and performance.

More specifically, the experiment addressed the following experimental questions:

1. How does time allocation (as measured by time spent on text, media, and task performance) differ as a function of experimental condition (graphics versus video)?
2. How does performance (as measured by quality, quantity, and ease-of-navigation) differ as a function of experimental condition?
3. How do users’ subjective-ratings of the effectiveness and usability of the learning system differ as a function of experimental condition?
4. What is the relationship between degree of previous experiences and time allocation?
5. What is the relationship between degree of previous experiences and performance?

Once again, a detailed quantitative and qualitative usability analysis was conducted. With respect to the first experimental question, there was very little difference between the two experimental conditions in terms of the amount of time they spent on the different types of scaffolds. However, it was interesting to note that, in general students spent most of their time reading the text and carrying out the task, as opposed to using either the video or graphics. With respect to the second and third experimental questions, which involved the degree to which the different scaffolds differed, the bulk of the
evidence indicated that those who had the video as a supplemental scaffold performed better than those who had the graphic screen shots. This is particularly interesting, given that the scaffold of choice (the text) was the most abstract with respect to exactly what the user was to do. Perhaps this indicates that, if a user needs to rely on additional media, more rich and representative media is preferred.

The fourth and fifth question addressed the impact of previous experience on time allocation and performance. A reasonably strong positive relationship was found between experience with Dreamweaver® and time spent on the task, while a strong negative relationship was found between Dreamweaver® experience and the time spent reading the text directions. Not surprisingly, as experience ratings increased, users tended to focus more on the task and less time reading directions. Also, consistent with expectations, across all measures, those with more experience performed better.

1.5. RESEARCH GOALS
This research will endeavor to:

♦ Evaluate a learning system for teaching civil engineers Geographic Information Systems in order to optimize its pedagogical effectiveness.

♦ Explore the efficacy of iterative usability testing as a tool for developing learning technologies
2. METHOD AND RESULTS

2.1. IDENTIFICATION AND TESTING OF AN EXISTING SYSTEM

2.1.1. Overview  After establishing some initial design parameters, a review of existing systems was carried out with the goal to identify a learning system that was close enough in design, to provide useful and meaningful data to guide future development and iterative testing of prototypes of the proposed system. In this case the most important requirements for such an existing system were: a) The training tasks were similar in that the learner was required to learn a software tool, within the context of specific closed-ended problems, and the learning system was external to the target software to be learned; b) The interface was similar to the interface design proposed for the GIS system, which was based on past research and guidelines associated with design of learning technologies; c) The system was web based, so that the technical components could, hypothetically, be translated into SCORM compliance; d) The content could be easily conceived as learning objects; and e) The system was designed to support multiple levels of scaffolding consistent with a progressive scaffolding model. The system should be similar in design constraints to the learner system, but does not necessarily have the overarching learning objectives associated. Given the time constraints of the project, the novel idea of furthering design from existing systems was readily accepted. Even the time required to make some low-fidelity prototypes could have been prohibitive towards the design cycle.

Microsoft’s E-Learning Library (MELL) was identified as a software system that met these criteria (http://www.microsoft.com/learning/mell/) Also see Figure 2.1 for a screen shot from the MELL system. The main goal of the Microsoft software is to educate end-users on how to use different Microsoft products in real-world situations. The training goals map well with those of our project, which is intended to instruct civil engineering students on how to use GIS software while reinforcing and applying skills learned in the traditional class. Both the GIS and Microsoft learning systems were designed to teach users specific software applications, using a separate learning system, within the context of specific tasks. The tasks of the software are different in the spectrum of training versus educating. The GIS project intended to produce software that
would educate and reinforce geotechnical concepts to users, while the MELL system has no counterpart. The MELL system strives for training users how to use Microsoft products.

Figure 2.1 - MELL Website Text and Video Demonstration

Although not SCORM compliant, the MELL system was web-based, which is an important and specific criterion for SCORM compliancy. The MELL system uses a proprietary back-end instead of the open standard produced by ADL consortium. However, the web-based structure of MELL mimicked the look and feel of a potentially SCORM compliant learner system.
Furthermore, the MELL system fit well with the notion of progressive scaffolding, which was an important criterion for GIS learner system design. As discussed previously, the use of different levels or scaffolds presumably provides greater pedagogical utility. The MELL system included animations, which offered demonstrations for the learner to follow and/or mimic, in addition to plain text descriptions to aid the user. Step-by-step text instructions were given in addition to animated demonstrations, developed in Macromedia Flash©, coinciding with each step. This text was equivalent to the video demonstration in having the exact same material but different levels of media to display it to the user. One can consider each tier a different fidelity of media but assistance as well, with higher support coming from the more detailed demonstrations as compared to the less detailed sidebar text descriptions. In like manner, the proposed GIS system intended to use at least two scaffolds of support for the learner to garner the benefits of the progressive scaffolding model. Thus, we judged the two systems as similar enough to provide testing, and, most importantly, a useful starting point for iterative usability testing to guide future development of our system.

This is not implying, however, that the two systems had overlapping goals. While the two systems face similar design concerns, the main objectives still remain separate: MELL wants to educate users on how to use MS Access, our learner system strives to educate or reinforce geotechnical engineering concepts through a specific problem. This difference in goals does not provide completely parallel mapping between the systems, but this same difference is expected to a degree. Part of this project’s novelty is that this learning system should provide higher level learning versus training. In essence, this project exists because few systems strive to reach this higher order goal, which makes the identification of a system that has both similar design constraints and educational objectives nearly impossible.

It was decided that our system interface would be modeled on this system, dependent on the results of initial usability testing of the MELL interface, even with an understanding of its limitations. A MELL module that covers Microsoft Access©, a database implementation was selected for this experiment. Four students were recruited from an Information Systems class in an Information Science and Technology (IST) department where database concepts are taught and practiced, and were given a specific
problem to solve using Microsoft Access. This testing mapped well to our project, since the students in the IST course had received instruction in the basic theory of databases, similar to our target subjects who will have had instruction in geotechnical theory. In essence, the test subjects for MELL and the intended end-users of the learner prototype had the pre-requisite understanding of their respective domains, but were being asked to implement it in an unfamiliar manner. In both cases the purpose of the learning system is to teach the students to use a software tool (i.e., Access or ArcView GIS), within the context of specific problems.

2.1.2. MELL Testing

2.1.2.1 Participants The study consisted of four (4) undergraduate students (3 male, 1 female) enrolled in IST 141- Information Systems in the summer semester of 2004. After filling out experiment consent forms, the subjects were given a brief 5-point (disagree – agree) Likert-scale questionnaire with the following questions:

- I have a lot of experience working with databases
- I have a lot of experience working with Microsoft Access
- I have a lot of experience working with Forms in Microsoft Access

2.1.2.2 Directions Subjects were given the following directions upon completion of the entrance questionnaire:

"You are being asked to use the Microsoft® E-Learning Library (MELL) (http://mell.umsystem.edu) as a reference for building forms within Microsoft Access.

You will be given a half-hour (30 min) to complete the allotted tasks. Finish as much of the tasks as possible within the time available, but completion is not a requirement. A video rental store database (Movie_Rental_Store.mdb) will be open on the testing computer, with both pre-constructed tables and records for the completion of a series of tasks. The MELL website will be open and set to the
correct page within the site. The experimenter will provide a brief familiarization of the system before you begin.

While you are working, please follow the “talk-aloud protocol,” in which you say what you are thinking while you are doing your work. The experimenter will keep track of the time and take notes. The experimenter will not be able to offer assistance with any of the tasks completion. The testing is over whenever you complete the required tasks or the half hour duration is exhausted. If you finish the tasks before the end of the 30 minutes, please inform the experimenter.

Please remember that the MELL website is a tool and does not necessarily need to be followed step by step- use it at your discretion.

Keep in mind that the intention of this study is to test the interface of this system, not you”.

The researcher then explained the highlights of the previously read directions and then gave the subject the chance to ask any questions.

The researcher took field notes while both the subject’s computer screen and their facial expressions were recorded during the research study.

2.1.2.3 Tasks Users were given the following three tasks of increasing difficulty to complete with the aid of the MELL website:

1. Build a form that contains/displays all information from the MOVIE table
2. Build a form that contains/displays all information from the MEMBERS table. Improve the appearance by changing the fields “mem_fname” and “mem_lname” to “First Name” and “Last Name” respectively.
3. Add a subform to task 2 form containing relevant information from the TRANSACTION table.

Example pictures were given for each task, but subjects were informed that mimicking the appearance was not required.
2.1.2.4 Post-testing questionnaire Once testing finished, a small post-testing questionnaire was administered. The user was asked to evaluate the following statements on a 5 point Likert-scale:

- I found it easy to complete the tasks from beginning to end
- The MELL website helped me greatly in performing the tasks I completed.
- The MELL website was easy to use.
- I relied a great deal on the website animations to complete the tasks
- I relied a great deal on the website text to complete the tasks

Table 2.1 contains the user responses. In addition to the Likert-scale questions above, users were also asked to provide any suggestions for improvement (see Section 2.2.5.1).

2.1.3. Results

2.1.3.1 Quantitative Results The following section contains the questionnaire responses as well as measures of performance. Qualitative notes and inferences are provided as well.
Table 2.1- Quantitative Testing Results of MELL

<table>
<thead>
<tr>
<th>Subject</th>
<th>Pr1</th>
<th>Pr2</th>
<th>Pr3</th>
<th>Po1</th>
<th>Po2</th>
<th>Po3</th>
<th>Po4</th>
<th>Po5</th>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
<th>Overall Time (mm:ss)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>complete</td>
<td>complete</td>
<td>complete</td>
<td>26:34</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>incomplete</td>
<td>incomplete</td>
<td>incomplete</td>
<td>18:08</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>complete</td>
<td>complete</td>
<td>complete</td>
<td>17:59</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>complete</td>
<td>complete</td>
<td>complete</td>
<td>19:33</td>
</tr>
<tr>
<td>Question Average</td>
<td>2.25</td>
<td>3</td>
<td>1.25</td>
<td>3</td>
<td>3</td>
<td>3.5</td>
<td>2.5</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SCALE: STRONGLY DISAGREE … 1 …. 2 …. 3 … 4 … 5 … STRONGLY AGREE

Entrance Questionnaire

Pr1: I have a lot of experience working with databases
Pr2: I have a lot of experience working with Microsoft Access
Pr3: I have a lot of experience working with Forms in Microsoft Access

Exit Questionnaire

SCALE: STRONGLY DISAGREE … 1 …. 2 …. 3 … 4 … 5 … STRONGLY AGREE

Po1: I found it easy to complete the tasks from beginning to end
Po2: The MELL website helped me greatly in performing the tasks I completed.
Po3: The MELL website was easy to use.
Po4: I relied a great deal on the website animations to complete the tasks
Po5: I relied a great deal on the website text to complete the tasks

2.1.3.2 Open Ended Questionnaire Responses

Recommendations from the users are listed below in response to the following open-ended question from the exit questionnaire:

What suggestions would you offer for improvement of the MELL interface so that it would make it easier for you to learn the materials?

Subject A

- Higher contrast mouse cursor on example animations
- Next step button slightly bigger and at top and bottom of window
Subject B

- It’s difficult to understand, make it easier maybe…

Subject C

- Possibly more descriptive text. I didn’t work with it enough to see if there were links to detailed descriptions for the tasks, but I think that would helped me, in an un-timed environment, get a firm grip on what it is I am actually doing.

Subject D

- I would not do much different … I would try to make the tutorials open in the same window and not in a new one (Frames)

2.1.3.3 Summary of observations from field notes The following section contains overall observations derived from observations and field notes categorized into major themes

1. Visual

- Video was too slow- Users seemed to find the video too slow and used it more after trying to complete a step based on the sidebar text of the Flash page, but doing so unsuccessfully. Users were able to directly follow the text a majority of time without need for higher scaffolds of assistance.
  - For example: Subject C was working on the first task and began watching one of the videos allocated for it. Subject C watched part of the first video, grew irritated and scanned through the video with the embedded flash navigation, then simply read the text instructions and worked on the target system without video afterwards.

- Video contrast could be improved -The visual contrast of certain actions within the videos was not high enough for certain users. Users D and C had to watch some animations a couple of times in
a row to see what was happening within them. Subject A provided a similar response in his exit questionnaire.

2. Scaffolding

- **Users begin by overestimating task difficulty** - Users started with the highest level of scaffolding available (video demonstrations) but moved towards lower scaffolds of help (from the sidebar text of the Flash page to none at all) as they grew comfortable with both the system and task list. Most users began the 1st task by getting some form of help from the MELL system—generally watching the video. In the 2nd and 3rd tasks, users began to wean off the MELL system for constant help and used it more to solve a specific challenge to completing the task at hand. Since the tasks were built of the previous task to some degree, the confidence and knowledge gained from completing the prior task gave the users momentum towards completing the next one.

- **Over-all participants relied on lower-level scaffolds** - Most users preferred to look at the step-by-step text accompanied by the still images versus the animation. The step by step hyperlinked text allowed the users to go to the sections of the demonstration that they had trouble with quickly. Moreover, most users found the text descriptive and detailed enough to complete the tasks without need of a demonstration.

3. Navigation

- **Linear navigation hidden from users** - Users initially had trouble finding the ‘next’ and ‘previous’ buttons on the animation page, which is particularly pragmatic due to the linear nature of the tasks they were assignment.

  o For example: Subject A had to search for “next” and “previous” buttons at the beginning of initial task. The subject began by looking for them at the top of task list. Even after finding the buttons on
the bottom of the task list, the user still searched for
the navigation buttons at the top of the task list.

2.1.4. Discussion Overall, users found the MELL system quite usable as exemplified by
75% of the users completing all required tasks. This is impressive when given that all but
one user rated themselves the as having no prior knowledge about forms within MS
Access. While Subject B was unable to complete any of the tasks, the problem did not
appear to be system oriented. However, the true cause of poor performance was unable
to be determined. While lack of motivation, energy, skills, and other explanations were
postulated, none could be proven beyond a doubt. The user was unable to be interviewed
which could have provided the insight required into why this situation occurred. Future
endeavors will strive to further investigate similar situations to more satisfying
completion. Both Subjects A and D found that the MELL website greatly helped their
completion of tasks as indicated in their exit questionnaire. Subject B seemed to find the
website useful despite being unable to complete the tasks with its aid. Subject C, who did
complete all the tasks successfully, did not feel that the website greatly helped the
completion of tasks. This correlates well with his pattern of use for the site, which was to
only use the site if he could not figure out the next step on his own first.

Given the over-all positive nature of the results, the team felt confident in moving
forward with the proposed interface design for the GIS system, which was similar to the
MELL interface. However, this initial testing also pointed to weaknesses that could be
addressed in development of the GIS prototype. These are listed in the specific
recommendations below.

The testing also provided useful insights into how the user went about using the
learning system in performing their tasks. In particular their use of the different
scaffolding media was instructive. Both the students’ questionnaire responses and
observation of their behavior indicated that they used the low level scaffolds to a greater
degree. Further, the reliance on low level scaffolds increased as they progressed through
the tasks, and the reliance on lower level scaffolds was greatest with those who were
most skilled. All of these general findings are consistent with the series of progressive
scaffolding studies discussed in the introduction (Hall, DiGennaro et al. 2002; Hall, Stark
et al. in press), which were carried out with learning systems designed to teach different content. This pattern of use is consistent with a user who recognizes the goal of performing specific tasks with the target software, while limiting interaction with the learning system to the minimum necessary to perform these tasks. In general the learning system is then truly a peripheral tool, which is only used when required, and only at the minimum level necessary for task performance. Therefore, in general, design practices that support this type of learning activity will be most effective.

2.1.5. Recommendations The following is a list of recommendations distilled from both observations and user suggestions which was provided to the development team for development of the original prototype:

- Model overall GIS interface after MELL interface
- Increase the contrast/size of the mouse cursor on the example animations
- Increase the size of the next step button and place it on both the top and bottom of the window
- Make the plain HTML page contain the same instructions as given as the sidebar on the animation page
- Support both static images as well as the dynamic animations (allow users to initiate playback for individual sub-tasks)

2.2. INITIAL PROTOTYPE USABILITY TESTING

2.2.1. Overview An initial prototype of the GIS learning system was developed, integrating the recommendations from the MELL assessment (Figure 2.2 & 2.3). Usability testing of this initial prototype was carried out in order to provide direction and improve the development of a GIS learner system that would be used within an ongoing class. This prototype represented the initial design choices for layout and only included a portion of the content to be incorporated in the production prototype.
Figure 2.2 – ArcView Learner System after Initial Prototype Recommendations Implemented

Figure 2.3 - Screenshot of video demonstrating ArcView fundamental after Initial Prototype Recommendations Implemented
2.2.2. **Participants & pre-questionnaire** This study consisted of four (4) undergraduate students (3 male, 1 female) who were enrolled (Summer 2004) in undergraduate civil engineering courses at the University of Missouri-Rolla. After filling out experiment consent forms, the subjects were given a brief questionnaire to determine their familiarity, enthusiasm and/or comfort level with ArcView GIS, geotechnical engineering. Subjects were also asked about the number of soils classes taken in the civil engineering department. The following is a list of the pre-questionnaire questions:

- I am familiar with ArcView GIS
- I am comfortable with ArcView GIS
- I am highly competent with ArcView GIS
- I am enthusiastic about GIS and Geotechnical engineering
- I have taken CE 215, Introduction to Soil Mechanics, (or equivalent)
- I have taken many soils classes

Table 2.2 contains the subjects’ 5-point Likert scale responses.

2.2.3. **Experiment design** Subjects were given the following directions upon completion of the entrance questionnaire:

“The intent of this study is to test a Learning System prototype developed for GIS.

You are asked to complete the modules in the Learning System given in sequential order. You will be given a half-hour (30 min) to complete the allotted tasks. Finish as much of the tasks as possible within the time available, but completion is not a requirement. The Learning System website will be open and set to the beginning page. Layers files are stored at C:\Documents and Settings\User\Desktop\Layers_Completed. The experimenter will provide you with a brief overview of the system before you begin.”
While you are working, the experimenter will keep track of the time and take notes. The experimenter will not be able to offer assistance with any of the tasks completion. The testing is over whenever you complete the required tasks or the half hour duration is exhausted. If you finish the tasks before the end of the 30 minutes, please inform the experimenter.

It is very important that you speak aloud what you are thinking while you go through the system. Even if you think what you are saying might sound dumb, this input can be invaluable in retooling the system.

Keep in mind that the intention of this study is to test the interface of this system, not you”.

The experimenter then reviewed the directions and answered any questions the participant had.

While the subjects proceeded with the given tasks, their computer screens, facial expressions were recorded and the experimenter took field notes concurrently.

2.2.4. Tasks The subjects were asked to watch and mimic all of the modules within the first section of the learning system, which was the entire prototype. The users were asked to go through in sequential order, but if a step could not be resolved, they were asked to skip it and move on to the next. At the end of the session, users should have completed a detailed and correctly formatted map.

This situation deviates from the planned lab deployment, in that the learner is not required to use the system, rather it is just to serve as support for their task carried out with the ArcView software. The view of the team was that the additional information to be gained from this thorough exploration of the learning system would outweigh information lost due to a less realistic task.
2.2.5. **Post-testing questionnaire** Once testing was completed, a small post-testing questionnaire was administered. The user was asked to evaluate the following statements on a 5-point (disagree-agree) Likert-scale:

- *I found it easy to complete the tasks from beginning to end.*
- *The provided website helped me greatly in performing the tasks I completed.*
- *The website was easy to use.*
- *I relied a great deal on the website animations to complete the tasks.*
- *I relied a great deal on the website text to complete the tasks.*

Table 2.2 contains the user responses.

In addition to the Likert-scale questions above, users were also asked to provide any suggestions for improvement.

*What suggestions would you offer for improvement of the website interface so that it would make it easier for you to learn the materials?*

2.2.6. **Results**

2.2.6.1 **Questionnaire and task completion tables** The following is the results from the entrance/exit questionnaires with task completion and time as well.


Table 2.2 - Quantitative Testing Results for Initial Prototype

<table>
<thead>
<tr>
<th>Subject</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
<th>Task In/Complete</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>No</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>complete</td>
<td>23:00</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>Yes</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>complete</td>
<td>26:32</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>Yes</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>complete</td>
<td>27:09</td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>No</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>complete</td>
<td>28:05</td>
</tr>
<tr>
<td>Q. Average</td>
<td>4.3</td>
<td>5</td>
<td>5</td>
<td>2.75</td>
<td>N/A</td>
<td>4.75</td>
<td>1.8</td>
<td>2</td>
<td>2</td>
<td>1.75</td>
<td>3.25</td>
<td>26:11</td>
<td></td>
</tr>
</tbody>
</table>

**SCALE: STRONGLY AGREE…1…2…3…4…5…STRONGLY DISAGREE**

**Entrance Questionnaire**

Q1: I am familiar with ArcView GIS
Q2: I am comfortable with ArcView GIS
Q3: I am highly competent with ArcView GIS
Q4: I am enthusiastic about GIS and Geotechnical engineering
Q5: I have taken CE 215, Introduction to Soil Mechanics, (or equivalent) *(Yes or No)*
Q6: I have taken many soils classes

**Exit Questionnaire**

Q7: I found it easy to complete the tasks from beginning to end.
Q8: The provided website helped me greatly in performing the tasks I completed.
Q9: The website was easy to use.
Q10: I relied a great deal on the website animations to complete the tasks.
Q11: I relied a great deal on the website text to complete the tasks.

**2.2.6.2 Open ended questionnaire responses** Recommendations from the users are listed below in response to the following open-ended question from the exit questionnaire:

*What suggestions would you offer for improvement of the website interface so that it would make it easier for you to learn the materials?*

**Subject A**

- Wanted some way to show one’s current location on the sidebar, or “some way of keeping track of what you should click next.”
- Requested bigger text.

**Subject B**

- If the website could suggest to the user to adjust the size of the windows so the user could see both the website/movie and the GIS
• Possibly make the website movies continue on after the user clicks a button on the movie (so it knows it can proceed).

Subject C
• Better Introduction
• Clickable (skip-able) control of video

Subject D
• I found the site very helpful and I personally can’t think of anything that would make it easier

2.2.6.3 Overall observations as recorded in field notes
The following observations were recorded during the testing sessions.

1. Visual & Audio
Video speed was a constant problem. All subjects found the video playback speed too slow at one time. Subject C was so disenchanted with watching the video that he resorted to skipping through the still images with questionable success. He found himself getting lost at times and having to repeatedly go back to certain areas of video. Subjects A and B also found the video too fast at times and had to re-watch certain areas that went too quickly. This was more a problem for subject B, however, since she was trying to work simultaneously with the learning system and the ArcView software. Subject B noted that she felt that the slower video was a better alternative to the fast since “at least you can skip ahead.”
  o Example: Subject B worked on the target system while watching the videos concurrently, and found that the video speed was less compatible with this learning style. When watching “View from top to bottom” video, subject B said “The movie almost moves...too fast. ... I can’t see how to get it [the task to be completed] fast enough” while trying to work simultaneously in the learner and target systems.

Users had trouble with the video’s inconsistent dialogue boxes. While watching video, Subject A declared that the “dialogue boxes are handy” in determining
what action is truly occurring on the screen. Further into the prototype, a new menu dialogue opened without such a box to which he stated, “I have to assume that I’m going to right click,” revealing some frustration. He further indicated his position by saying “It’s (the video) kind of real detailed on some things and not on others.” Right-click dialogue boxes were the most absent dialogues.

Audio feedback was helpful. Subjects were assisted by the audible feedback provided whenever an action (clicks) occurred. Subject B specifically mentioned this being a helpful function. She didn’t notice the visual feedback of the cursor flashing and suggested that it did flash (more).

Multiple Windows were Problematic. Participants as a whole were bothered by each video opening in a separate new window.
  
  Example: Subject C got lost with the sheer number of video windows open, and had to close them all up and re-open the requested video to assure himself he was watching the correct video.

2. Navigation

Navigation between learning module prototype and ArcView GIS
Subjects A, B, and D followed the videos directly, with subjects A and D alternating between watching video and then completing the exercise, while subject B worked simultaneously by watching the video and working in ArcView concurrently.

Navigation within the learning module prototype
Subject C worked through the exercises on a by need basis, skipping through the video to interested parts.

Difficulties with Navigation Tree
All subjects had some level of trouble with the navigational tree, from looking at different sections to determine next section to skipping or re-watching a section.
Little use of Linear Navigation Support

Back and forward buttons currently implemented were not used and seemed to be out of the subjects view.

- Example: SCORM interface provides linear navigation, but it was out of users’ locus of attention. After subject A skipped a step early on, he explained a preference for a more linear style. The linear style navigation was, however, supplied but unnoticed by the participant. He mentioned the navigation tree did not fit his navigation style: “Having all of these available, even for smart, intelligent people like engineers, maybe isn’t the best idea.” He further described a linear navigation he would prefer, despite such a navigation already existing: “They’d start here [subject pointed to the first position of the navigation tree]... but you like click a ‘next’ button.... maybe a ‘next’ and ‘previous’ buttons.”

3. Content

All subjects felt more of an introduction could be useful. Even after having the read the instructions and been briefed again on them, they were not comfortable with what was needed to be completed. Subject B noted that she felt “coming in cold” was hard while subject C wanted a picture of the final product at the start page to smooth the transition.

2.2.7. Discussion

In terms of overall effectiveness of the interface design, the results were generally positive, indicative of a design that effectively supported user activities. All users completed the objective to produce the detailed map in ArcView GIS, but with varying levels of quality. The fact that all users completed the task is still noteworthy, given that the pre-questionnaire indicated that no users felt either comfortable or confident with ArcView GIS software (average user Likert- score on both measures: 5 or strongly disagree with the statement). The post questionnaire responses were consistent with the high completion rate, in that users found the completion of tasks easy when facilitated with the learning module (mean score of 1.8 for Q7 and 2 for Q8). The website and its design also rated highly among the participants with a mean score of 2 for
Q9. The student’s use of the various scaffolds, again, provided interesting information about their use of the tool. Users unanimously preferred the high scaffold of animation compared to lower scaffold of text as shown through their likert-scale responses for questions Q10 and Q11 with mean scores of 1.75 and 3.25 respectively. These findings are seemingly inconsistent with the results of the MELL testing, which was interpreted as indicating that users tended to utilize the minimal scaffold necessary to perform the task. The problem noted with the lower scaffold of text was that it was not truly developed in this prototype. Users were forced to use the higher scaffold of video, which was complete for the tasks required. However, even with the necessity of videos being used, these findings still provided more detailed insight into issues that were associated with the tendency to avoid or become frustrated with the higher-level scaffolds. First, participants had issues with the speed of video playback across the board. When watching videos that showed rather simplistic tasks, users became bored and frustrated with a video that was too elementary. On the other hand, when working on a complex problem and trying to follow the accompanying video, users became flustered with video whose speed exceeded their ability to work. Instead of having video that is tailored towards either end of the speed-complexity spectrum, fast for simple tasks or slow for complex ones, the video speed is set at an in-between setting, which frustrates users in both complex and simple tasks. This points to the need for control over the video. By having the end-user in charge of the video playback, any adjustments required can be made while working, minimizing user frustration. Furthermore, the detail of the video was not consistent: at times it contained high levels and other times strikingly less detail. This frustrated users to some degree- users felt uncomfortable about having to make assumptions and should not be forced to.

The issues associated with the use of multiple scaffolding levels, in particular the video, lead to some more general design principles, which are again consistent with the view of a user who is using the learning system as a flexible tool when necessary, while focusing on a specific task. First, it is important to provide the user with as much control as possible, particularly with the higher level scaffolds, to provide the flexibility necessary for her to focus on specific needs, without spending an inordinate amount of time or effort. Second, it’s important to provide consistency, as exemplified by user
frustration with the inconsistent level of detail. Consistency allows the learner to more readily form a mental model, which represents the whole system, facilitating quicker scanning and search for specific information necessary to perform tasks.

2.2.8. Recommendations The following is a list of recommendations distilled from both observations and user suggestions:

1. Review dialogue prompts and strive for consistency
   - Sometimes high levels of detail, other times low levels
   - Right-Clicks weren’t always displayed. Clicking “Ok” always got a dialogue
   - The same user who commented that “These dialogue buttons are handy” also said “I have to assume I’m going to right-click here…”

2. Add static text on how to complete the tasks
   - Would help users who find the video too slow
   - The user could discover whether watching the video will answer the learner’s question(s)

3. Make an overall “Introduction” section to the lab
   - Even after being told that the purpose of this section of exercises was to produce a map in GIS, users still felt a bit lost

4. Make movies open in same window
   - Users found it difficult to select the right video when more than video web-page was open

5. Add a title to video web-page that correlates with the section heading
   - Help avoid user confusion by offering

6. Improve video navigation
   - Add sub-tasks with markers
   - Fast Forward that doesn’t cut out frames, but cuts the amount of time a frame is displayed
   - One user didn’t watch a video through after the first one- used flash navigation bar instead. Missed/poorly accomplished some steps
   - Another user who watched videos entirely did not quite finish tasks
7. **Increase visual contrast of cursor for added feedback**
   - Added contrast will reduce videos
   - No audio feedback provided in civil engineering computing lab- Visual feedback should compensate

8. **Add redundant next and previous links in the main frame**
   - Users sometimes got lost even with the matching title and tree heading
   - Previous and next links were not discovered by users looking for them

### 2.3. FINAL PROTOTYPE USABILITY TESTING

The usability of the production ArcView GIS Learner System prototype (Figure 3) was tested in order to provide recommendations for final improvements prior to implementation in a classroom setting.

#### 2.3.1. Participants & pre-questionnaire

This study consisted of five (5) senior undergraduate students (5 male) who were currently enrolled (Fall 2004) in civil engineering at University of Missouri-Rolla. Three (3) of the participants worked individually while the remaining two (2) worked cooperatively. After filling out experiment consent forms, the subjects were given a brief questionnaire to determine their familiarity, enthusiasm and/or comfort level with ArcView GIS, geotechnical engineering. Subjects were also asked about the number of soils classes taken in the civil engineering department. The following is a list of the pre-questionnaire questions:

*I am familiar with ArcView GIS*
*I am comfortable with ArcView GIS*
*I am highly competent with ArcView GIS*
*I am enthusiastic about GIS and Geotechnical engineering*
*I have taken CE 215, Introduction to Soil Mechanics, (or equivalent)*
*I have taken many soils classes*

#### 2.3.2. Experiment design

Subjects were given the following directions after the completion of the entrance questionnaire:
The intent of this study is to test a Learning System prototype developed for GIS.

You are asked to complete the CE 215 Soils lab 6. You will be given two (2) hours to complete the allotted tasks. Finish as much of the tasks as possible within the time available. The Learning System website will be open and set to the beginning page. GIS layers files are stored at C:\Documents and Settings\User\Desktop\GIS_Layers. The experimenter will provide you with a brief overview of the system before you begin.

While you are working, the experimenter will keep track of the time and take notes. The experimenter will not be able to offer assistance with any of the tasks completion. The testing is over whenever you complete the required tasks or the two hour duration is exhausted. If you finish the tasks before the end of the two hours, please inform the experimenter.

It is very important that you speak aloud what you are thinking while you go through the system. If something is not where you expected it to be or works in a different manner than you would perceive, let us know. Even if you think what you are saying might sound dumb, this input can be invaluable in retooling the system.

Keep in mind that the intention of this study is to test the interface of this system, not you.

The experimenter then reviewed the directions and answered any questions the participant might have about the directions.

2.3.3. Tasks Subjects were then presented with the following lab handout, which outlines the tasks required, and requested deliverables:
Some construction projects with significant earthwork operations required importing soil from a borrow source. Importing soil means that the soil will be obtained from a borrow source outside of the project boundaries. To select the borrow site we need to define what are the material requirements, which depends on the engineering objectives.

The engineering objectives of the earthwork construction are defined in the design phase of a project. For example, if a landfill is being built with an impermeable liner as the bottom layer, then a compacted clay soil is the material requirement. For the landfill example, the engineering objective is an impermeable liner and the material requirement is a compacted clay layer. In this laboratory you will be using a Geographic Information System (GIS) to solve an engineering problem. A GIS is system composed of electronic maps, databases and software tools. A software package manages this information and allows you to perform analysis to support engineering decisions.

Your group will be assigned a construction site with a particular engineering objective. You are to select one soil borrow site for the construction site you were assigned. In addition to meeting the engineering objective your selection needs to be the most cost effective.

<table>
<thead>
<tr>
<th>Construction Site</th>
<th>Engineering Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>St. Louis Stadium</td>
<td>Structural Fill</td>
</tr>
<tr>
<td>Fenton Landfill</td>
<td>Landfill Liner</td>
</tr>
<tr>
<td>Chesterfield Bottoms</td>
<td>Subsurface Drain</td>
</tr>
</tbody>
</table>
LAB OBJECTIVES:
Upon completion of this lab you should be able to:

1. Define what are the engineering objectives and material requirements for a construction earthwork operation.
2. Select an appropriate soil borrow site for a particular construction site.
3. Use a Geographic Information System for the selection of a borrow site.

EQUIPMENT/MATERIALS:

- Computer in CLC Rm. 115
- Software - ArcGIS®
- Data Packet (installed at C:/ GIS_Layers)

PROCEDURE:
A web-based learning system has been developed to guide you through the procedure on how to explore, examine and analyze the spatial data to support your decision on the selection of the appropriate and most cost effective soil borrow site for the construction site assigned to you.

DELIVERABLES -- TO DO and TURN IN:

1. Statement with the name of the soil borrow site selected.
2. Map printout showing the geology OR soils of the borrow site with roadways and construction site shown.
3. List of laboratory tests used to determine the soil type (USCS symbol) at the borrow site. Include the laboratory test results obtained from the testing lab (this is the email you received from lab with invoice).
4. Cost of the imported soil including trucking costs.
5. Justification statement of why this selected soil borrow site is recommended. This is a paragraph (letter) that you will write to a client explaining your recommendation.

Subjects were informed that the deliverables were due at the end of the two hour period at the latest. For deliverable #2, subjects were allowed to merely produce a map in ArcView and not forced to make a printout. Subjects were free to use the Learning
Module as they saw fit. However, when the experimenter was asked questions during testing, participants were generally directed back to the learner system for answers.

2.3.4. Post-testing questionnaire Once testing was completed, a post-testing questionnaire was administered. The questionnaire was the same questionnaire to be administered after the quantitative study, and thus contained some questions which relate to the lab sections and class in general but are irrelevant in this testing. While some questions were irrelevant to this study, the questionnaire was left un-edited per request of the professor in charge of the class in order to get an estimate of the time required to complete the questionnaire. The first two pages of the questionnaire consisted of the following questions:

Please use the scale below to respond to each of the statements and explain your answers in the space following, if appropriate.

Strongly Disagree 1 … 2 … 3 … 4 … 5 … 6 … 7 … 8 … 9 Strongly Agree

_____ 1. I learned a great deal of information about soil borrow site selection from this week’s lab.
_____ 2. I learned a great deal of information about soil borrow site selection from class lectures.
_____ 3. I learned a great deal of information about soil borrow site selection from the class text.
_____ 4. I found this week’s lab on soil borrow site selection to be very motivational.
_____ 5. I found the class lectures over soil borrow site selection to be very motivational.
_____ 6. I found the class textbook’s coverage of soil borrow site selection to be very motivational.
_____ 7. This week’s lab activity over soil borrow sites was applicable to “real world” engineering.
_____ 8. The class lecture over soil borrow sites was applicable to “real world” engineering.
_____ 9. The textbook coverage of soil borrow sites was applicable to “real world” engineering.
_____ 10. Before the lab activity that covered soil borrow sites, I knew a great deal about the subject.
_____ 11. After the lab activity that covered soil borrow sites, I knew a great deal about the subject.
12. Please list the strengths of the lab activity that covered soil borrow sites, in terms of it’s effect on learning and motivation, and it’s applicability to “real world” engineering.

13. Please list ways in which the lab activity that covered soil borrow sites could be improved.

After each question, students were left with an area to further explain their answers and given a prompt to do so. An extended amount of room was left for the final two questions. The third page consisted solely of technical questions. A list of these questions can be found in Appendix D.

2.3.5. Quantitative Results The following section contains the questionnaire responses as well as measures of performance. Questions participants found irrelevant (questions aimed at students in the experimental class situation) were excluded from this table. However, question numbers remain consistent with the original question set provided.
Table 2.3 – Final Prototype Testing Results

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Pr1</th>
<th>Pr2</th>
<th>Pr3</th>
<th>Pr4</th>
<th>Pr5</th>
<th>Pr6</th>
<th>Po1</th>
<th>Po4</th>
<th>Po7</th>
<th>Po10</th>
<th>Po11</th>
<th>To</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>yes</td>
<td>2</td>
<td>9</td>
<td>7</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>83.3</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>yes</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>91.7</td>
</tr>
<tr>
<td>C1</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>yes</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>83.3</td>
</tr>
<tr>
<td>C2</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>yes</td>
<td>3</td>
<td>7</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>83.3</td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>yes</td>
<td>2</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>91.7</td>
</tr>
<tr>
<td>Question</td>
<td>5</td>
<td>4.8</td>
<td>5</td>
<td>2.2</td>
<td>N/A</td>
<td>2.6</td>
<td>7.4</td>
<td>7</td>
<td>8</td>
<td>6.8</td>
<td>8</td>
<td>86.7</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pr1  I am familiar with ArcView GIS
Pr2  I am comfortable with ArcView GIS
Pr3  I am highly competent with ArcView GIS
Pr4  I am enthusiastic about GIS and Geotechnical engineering
Pr5  I have taken CE 215, Introduction to Soil Mechanics, (or equivalent)
Pr6  I have taken many soils classes

Scale: Strongly Agree 1 … 2 … 3 … 4 … 5 Strongly disagree

Po1  I learned a great deal of information about soil borrow site selection from this week’s lab.
Po4  I found this week’s lab on soil borrow site selection to be very motivational.
Po7  This week’s lab activity over soil borrow sites was applicable to “real world” engineering.
Po10 Before the lab activity that covered soil borrow sites, I knew a great deal about the subject.
Po11 After the lab activity that covered soil borrow sites, I knew a great deal about the subject.

Scale: Strongly Disagree 1 … 2 … 3 … 4 … 5 … 6 … 7 … 8 … 9 Strongly Agree

To  Percentage of technical oriented questions answered correctly (see Appendix D)

2.3.6. Qualitative Results

2.3.6.1 Subject Recommendations from Exit Questionnaire  Recommendations from the users are listed below in response to the following open-ended question from the exit questionnaire:

*Please list ways in which the lab activity that covered soil borrow sites could be improved.*

**Subject A**

- Information on which type of soil you want for the foundation and unit weight for all [could be provided]

**Subject B**
• Make the instruction[s] a little more clear in certain areas.

Subject C1
• I think the units where it said the truck carried 15 tons was supposed to be 15 CY, because we needed 880 CY and didn’t know how much CY [cubic yards] a truck carried. The very last section on “Proposed problem” seemed like it could have been up further in the outline, because we needed that to make our choice on what borrow site to choose.

Subject C2
• Some of the steps seemed redundant. It wasn’t always easy to find the information required to solve the problems. Maybe get rid of some of the extra layers.

Subject D
• None come to mind.

2.3.6.2 Overall observations as recorded in field notes The following section categorizes the observations garnered from the notes and videos.

1. Visual

Despite finding the video was helpful, “Reading is alright, but showing something is better” – subject A, videos were not watched often and were considered slow.

2. Navigation

The navigation section addresses two distinct types of navigational interaction: Navigation within the learning module prototype and navigation between the learning module prototype and the ArcView application.

Navigation within learning module
• Group C went further into the module to determine the specifics of the lab objectives and deliverables, in spite of the fact that the material was presented in a hard copy earlier. This group used the tree structure exclusively.
• Subject D used the embedded “next” and “previous” buttons on the flat HTML pages implemented from earlier recommendations exclusively for navigation within the system.

• No subject or group became lost within the learning module hierarchy while working on the lab.

Navigation between learning module prototype and Arcview application

• Users who switched back and forth between a fully expanded prototype and Arcview application window became irritated relatively early with this configuration. Two subject groups interacted with the two systems this way.
  
  o For example: Subject A sighed loudly while switching between a full-screen learner system and the target application when working on “Setting layer color schemes” after having done so two times before.

• Users/Groups who attempted to fit both the application and prototype to be visible simultaneously had an easier time following the directions as compared to when they worked between the two systems. Two subject groups worked in this fashion.
  
  o For example: Group C switched to this scheme early (4 minutes into their test) as did subject D, who would read the learning systems text as he worked in the system. After subject D switched interaction styles, he had to go back and forth between the systems less in the next task, as gauged by cursor and eye movement between the systems.

3. Scaffolding

• Users who started with the higher scaffold of video quickly went to just using text to accomplish tasks.
  
  o For example: Subject A began by recreating steps from the videos, but after 18 minutes, videos were only used as a second
resort. The subject would read the text supplied for the majority of assistance afterwards, while at times s/he tried work without assistance of the learner system beyond reading the title of the next step.

- Video was used as a clarifier when text was insufficient to accomplish a required task.
- Users would often re-read text first before moving to the higher scaffold of video.
- Group C, which consisted of two individuals, worked out confusions about the text descriptions between each other, which in effect served as an additional scaffold.
- Users tended to dislike the longer text descriptions and experienced more trouble with them.

4. Content

Text clarification

- Many users/groups mistakenly turned off a toolbar in step “I.e.i.1. Opening the tools toolbar” which was intended to turn that toolbar on (the toolbar was on by default). The text did not mention the intent of steps.
  - Subject A followed the instructions, and after seeing the intended results were not displayed (the toolbar was not available), s/he went back to the text instructions to see if the steps were followed correctly. After determining that s/he did follow the instructions correctly, stated “Where’d that go?”

- Users were confused by section “IV.b.i Segment Distances” instruction to double click distance tool since it requires a single click function.

- The video for “V.f.vi.1. Adding text callouts” doesn’t mention to hold the left click button and resulted in users having to reiterate this step.
2.3.7. **Discussion** This final iteration of testing revealed a prototype that appeared to be substantially improved from the original, with all users completing the required tasks as much as technically possible (the webpage required for “soil testing” results was down during some users testing and they had make decisions without that information). Not only were all objectives completed by the users within the two-hour time block, but they deemed the experience as generally favorable and improvements focused around content clarifications. The quantitative results added further support to the finding that this production prototype was generally effective. Users felt that they had learned a great deal about borrow site selection from this lab, with an average user response of 7.4 out of 9. In regards to its motivational factor, they again rated it above average with a mean score of 7, implying this lab is more motivating than the average civil engineering lab. They indicated that the activity was highly indicative of “real world” engineering, with an average score of 8 out of 9. Overall, they felt that their knowledge on borrow sites increased rather dramatically from the lab with mean before and after the lab scores of 6.8 and 8 out of 9 respectively. Qualitatively the prototype implementation proved to be a success as well. Users were able to make good use of the tool provided, and minimal questions were asked of the researcher during the study. Users were able to adjust to the ArcView application quickly and transfer skills/functionality learned in an early module section to accomplish a task required for later sections. It’s also important to note that, despite the participants all being accomplished civil engineering students (all were seniors) with some expertise in soils and geotechnical engineering, they all had little or no experience with ArcView GIS or any other GIS program. They had all taken Civil Engineering 215, the intended venue for the prototype deployment.

The navigation within the learner module also appeared sufficient based on users’ behavior, in that none got lost, and based on self-report, in that users did indicate that this was a problem. It was interesting to note that there was substantial variance in the manner in which navigation was used, with some relying most heavily on the tree structure and others relying on the linear navigation. This flexibility built into the system afforded them the opportunity to fit the system to their mental models. On the other hand, learners did encounter some difficulties with navigating between the learning system and the target software (ArcView GIS). This was true despite the fact that all
groups were informed that they were free to re-arrange the window configuration, in addition to their physical configuration, as they saw fit. One reason that they might have chosen to switch between the two full size windows was simply that the smaller window size was not adequate.

Despite the fact, that the videos had been changed based on recommendations from the initial prototype testing, all users still had some level of discomfort with watching the videos. The control provided was still insufficient which made for movies that were hard to speed up or slow down.

The learners’ use of various scaffolds went smoothly with users finding the level of scaffold necessary with minimal effort. As with the previous testing, with the exception of the under-developed initial prototype and its text scaffold, users preferred the more minimal level of scaffolds, and the process of users beginning with minimal scaffolds and moving to a higher level if the minimal failed, was even more evident in this research.

This test was the only test in which a dyad was tested, where two users performed the task together. Due to the fact that there was only one group tested, it’s difficult to generalize much from the results, but it is interesting to note that the presence of a second user served, in a sense, as an additional scaffold. When a problem was encountered after viewing the text directions, the partners would consult one another before utilizing the higher level video scaffold.

2.3.8. Recommendations The following is a list of recommendations derived from both observations and user suggestions which was provided to the development team for final modifications before going live:

1. Opening the tools toolbar

Users turned the toolbar off inadvertently by following the directions and then looked for the tools toolbar in vain. Others asked if they should turn it off.

“\textbf{\textgreater\textgreater} Left Click \textit{“Tools” from the drop-down menu}” should be appended with “\textit{if not currently selected}.”
2. **Segment Distances**

   “Double Click the “Measure” tool...” should be replaced with “Single click the “Measure” tool...”

3. **Adding text callouts**

   Should mention to click and hold callout button for placement.

   “Click and hold mouse button on the point where the balloon will point to >> Release mouse button on the point where the balloon will be placed”

4. **Determining cost of the project**

   Users were unsure how to estimate the tons/unit volume- perhaps further explanation could be given.

   \[(\text{Needed volume}) \times (\text{Estimated tons/Unit volume}) = \text{Tons of material needed}\]

**Future Implementation Recommendations**

1. **Dynamic Videos**

   Videos overall were deemed to be slow, but this alternated between a disadvantage and an advantage given the task.

   Offering users more control of the video would allow them to use a “custom fit” approach on how to respond to a given task.
3. DISCUSSION & CONCLUSION

3.1. OVERVIEW
Iterative usability testing proved to be an effective tool for developing learning technologies. Results of the usability testing indicated that the iterations were systematically improved based on the results and recommendations derived from usability testing. The discussion that follows is based on the research objectives posed earlier:

♦ Evaluate a learning system for teaching civil engineers Geographic Information Systems in order to optimize its pedagogical effectiveness.

♦ Explore the efficacy of iterative usability testing as a tool for developing learning technologies.

3.2. OBJECTIVE 1: EVALUATING MODULE FOR PEDAGOGICAL EFFECTIVENESS

3.2.1. Overview The discussion of the first objective will begin with a focus on the lessons learned that directly informed development, in the form of recommendations. This will be presented as a progression through the three rounds of usability testing, in order to illustrate the progressive and iterative nature of the process. This will be followed by a description of general insights gained regarding the way in which users tended to work with a learning system such as this, and general design principles based on these insights.

3.2.2. Usability testing of an existing system The MELL system was tested first in an attempt to investigate a system with design and functionality similar to the proposed GIS prototype. The four undergraduates who participated provided us with useful insights into how end-users interact with and manipulate a learning technology towards accomplishing learning goals as well as the strengths and weaknesses of the MELL design. Three of the four participants were able to make good use of this prototype with most users completing all required tasks with a high degree of competency. The students made use of the multiple layers of assistance. Text based help was rated to offer more assistance than its video counterpart, but the video scaffold allowed users to acquire more
in-depth help when text proved insufficient, and also served as a tool for checking their work. This initial evaluation of an analogous system also indicated the need for some design modifications, including issues with video and navigational components such as video playback speed, visual contrast of actions, and navigational tools being out of sight.

Recommendations were made to not only improve upon these deficiencies but also to emphasize the need to implement the system strengths. The most general recommendation, in terms of design, was to proceed with plans to develop an interface for the GIS system similar to the MELL interface. While the general interface design would remain the same, to make the first GIS prototype more usable, the following recommendations were proposed for development of the initial prototype:

- **Increase the contrast/size of the mouse cursor on the example Flash animations**
- **Increase the size of the next step button and place it on both the top and bottom of the window**
- **Make the plain HTML page contain the same instructions as given as the sidebar on the Flash page**
- **Allow users to initiate playback for individual sub-tasks in the video**

### 3.2.3. Usability testing of initial prototype

The first round of testing for the project teams’ initial design, again, consisted of four undergraduates, but since this prototype was developed with target population in mind, civil engineering undergraduates were excellent candidates to test the system. With the intended audience testing the system, new insights were garnered. First, video speed was again found to be a problem, with users complaining that the videos were slower than they needed to be, once again taking time away from their opportunity to interact with the target software, Second, dialogue boxes, which were added to the videos, to provide users with additional information, were inconsistent in the amount of information they contained, some times extremely detailed and at other times very general – this inconsistency clearly bothered the users. With regard to navigation, users had trouble with the tree structure provided and at times became lost in the system. While the SCORM interface did contain embedded next and previous buttons, most users did not use them. Finally, even after being given a general
overview of the tests purposes and the prototype, users still had trouble understanding what their true objective was.

Recommendations based on these findings were:

- **Review dialogue prompts and strive for consistency**
- **Add static text on how to complete the tasks**
- **Make an overall “Introduction” section to the lab**
- **Make movies open in same window**
- **Add a title to video web-page consistent with the section heading**
- **Improve video navigation**
- **Increase visual contrast of cursor for added feedback**
- **Add redundant next and previous links in the main frame**

Despite the large number of recommendations for improvement, over-all users found the system to be quite usable. The exit questionnaire results showed that, not only did they find the tasks to be relatively easy, but found the website easy to use as well.

### 3.2.4. Usability testing of production prototype

The prototype was again tested by four undergraduates in civil engineering after recommendations were implemented and the rest of the content was developed, but prior to a larger applied experiment to be carried out within the context of a Civil Engineering class. Though over-all the system performed much better than the previous iteration, issues in need of modification were identified. Video was again a culprit as users still found it too slow to aid the completion of all tasks. Navigation functioned much more effectively than previous iterations, presumably more consistent with users’ mental models. Users still began by over-estimating task difficulty and using the higher scaffold of video, but quickly reverted to the lower scaffold of text for the majority of steps. There were some frustrations in users when text descriptions became too long and they consequently had more trouble with those sections. Content disparities were noted as well, with an eye towards making a seamless prototype.

Most recommendations focused on content, indicating that most significant interface issues had been adequately addressed. The following is a list of recommendations distilled from both observations and user suggestions, provided to the
development team for final modifications of the production prototype to be used in the applied experiment:

- **Opening the tools toolbar**
  “>>& Left Click “Tools” from the drop-down menu” should be appended with “if not currently selected.”

- **Segment Distances**
  “Double Click the “Measure” tool…” should be replaced with “Single click the “Measure” tool…”

- **Adding text callouts**
  Should mention to click and hold callout button for placement.
  “Click and hold mouse button on the point where the balloon will point to >> Release mouse button on the point where the balloon will be placed”

- **Determining cost of the project**
  Users were unsure how to estimate the tons/unit volume- perhaps further explanation could be given.

- **Dynamic Videos**
  Videos overall were deemed to be slow, but this alternated between a disadvantage and an advantage given the task.

This iteration of development received the highest levels of user ratings. Quantitatively, the participants rated this system and the overall lab highly, with various scores indicating as such. Users felt that they had learned a great deal about borrow site selection from this lab, with an average user response of 7.4 out of 9. In regards to its motivational factor, they again rated it above average with a mean score of 7, implying this lab is more motivating than the average civil engineering lab. They indicated that the lab was highly indicative of “real world” engineering, with an average score of 8 out of 9. Overall, they felt that their knowledge on borrow sites increased rather dramatically from the lab with mean before and after the lab scores of 6.8 and 8 out of 9 respectively.

Qualitatively the prototype implementation proved to be a success as well. Users were able to make good use of the tool provided, and minimal questions were asked of the researcher during the study. Users were able to adjust to the ArcView application
quickly and transfer skills/functionality learned in an early module section to accomplish a task required for later sections.

The prototype navigation worked well because it did not dictate how a user would be forced to use the system. Users were allowed to select the navigational model that fit best with their personal mental models. This choice afforded users the ability to focus in on tasks and less on how to achieve the tasks through the navigation provided.

3.2.5. General Insights and Design Recommendations

Taken together, these results point to the importance of a learning tool as a flexible support system – the tool is not the task. This was evident even in the initial prototype study where learners were required to work with all components of the tool. Users in this study were bothered with having to watch slow videos when a text description was all they felt they needed. There is no need to require users to engage all parts of the learning system, when all they need to do is to accomplish their tasks. This is why the navigation of the final prototype had minimal problems during testing. The navigational scheme implemented allowed for more flexible navigation than previous iterations. Users were able to apply the navigational scheme that most facilitated task completion. Across all experiments learners tried to minimize their interaction with the tool by relying on minimal scaffolds as much as possible. Again, users do not want to spend more time on additional tasks than necessary. The lower scaffolds, despite being more abstract, offer the quickest interaction. While following a video might not leave much room for misinterpretations, there is a larger price in terms of time allotted that could be spent directly on task completion. Users are willing to accept that some mistakes will be made by using these lower scaffolds. Students were also most critical of, and performance was most hampered by, situations where they were unable to quickly locate the important information which would allow them to perform their primary task of learning the software. When users were not given the opportunity to work through the system deftly, their performance suffered and agitation arose. Not only do users like quick learner system interaction to be able to move to the text task, a lack of quick interaction encourages poorer performance.
This leads to at least three general design principles, which apply in systems such as that developed in this project. First, the presence of multiple scaffolds is helpful since it allows the user to select the level of scaffold that is needed at any given time. Second, it is important that the learning system be flexible, including multiple types of navigation schemes (i.e., hierarchical and linear), and allowing for easy access of the necessary scaffold at any given time. Third, it is particularly important that the content be structured consistently and logically, allowing the learner to quickly develop an effective mental model, since the user needs to be able to quickly identify and acquire specific content.

3.3. EXPLORE THE EFFICACY OF ITERATIVE USABILITY TESTING AS A TOOL FOR DEVELOPING LEARNING TECHNOLOGIES

This project yielded a number of important insights about the efficacy of iterative usability testing as a tool for enhancing the effectiveness of learning technologies. Overall, the approach did indeed appear to result in a more pedagogically sound and usable system. Which each step of the evaluation process the system improved based on observations of users and their feedback. The MELL website tested well and was acceptable to the users. The initial prototype, despite drawbacks and limitations, tested well also. The final iteration tested significantly better than the previous two. This growth of user acceptance is one indicator that usability testing integrated into the development process leads to more effective learning technologies. In addition to indicating that the method can be effective over-all, other more subtle issues that mediate this effectiveness were identified in this project. The two factors were: a) resource constraints and b) the need to educate team members on the role and importance of usability testing.

Resources were a scarce commodity on this project, with the scarcest being time. One of the project’s stated goals was to make the solution reusable through the use of SCORM-compliant technology, which helped narrow the specifications down to a certain degree, but this, of course, was only a small part of the design process. While ideally a multitude of configurations could be tested to determine which combination of user control, scaffolding levels, navigation scheme were ideal, restrictions on resources such as time, limited participant pool, and development manpower made such an idealistic solution impossible. Time for the project was already quite limited, with only one-year’s
duration to not only develop, but also implement, the intended solution. A compromise was accepted between the two contrasting factions of needs: the need for a highly usable learning tool and the need to remain within the resources allocated for project. This constraint was one important reason why the initial testing began with a fully developed analogous system, which allowed testing and development early in the project's life, without having to develop materials to test.

Educating team members on the role and importance of usability testing in the development process was also an obstacle to overcome, with many teammates having an unclear understanding of the role of usability testing within the development cycle. Many thought that usability testing was a form of bug or error checking based on checklists, and while bugs and errors may be discovered, this is not the intended goal. The goal of usability testing is to make systems and tools readily usable by individuals. Usability had to be “sold” throughout the project duration, with results and recommendations explained at various levels. Group dynamics also played an important role towards mediating the effectiveness of usability testing. Care and tact were necessary at every step. Designers grow attached to their designs to some degree. Usability testing puts one in an awkward position of informing a contemporary that their design is not the design preferred by the users. This needs to be done both skillfully and articulately, with explanations for every intended change. A good working relationship within the team is required to be able to pursue the highest levels of usability a product can afford.

3.4. FUTURE WORK

While the work and studies completed revealed many lessons, there are a number of issues still to be addressed with respect to the application of human-computer interaction methodologies with learning technology development. At the completion of thesis work, a larger scale applied evaluation of the learner system within the context of a class is being conducted. As mentioned previously, the results and insights provided from this experiment will offer a more applied evaluation of the system within the context where it was intended to be used. When these results are coupled and compared with the results from the usability testing experiments, a system for improvement can be synthesized through triangulation of findings from the usability and applied evaluations.
This synthesis will help ensure that future editions of this and other learner systems will be produced with high levels of usability and pedagogical effectiveness.

Future iterations can focus more on the development of multimedia with a greater degree of user control, particularly with the videos. When users have more control over the playback of the videos, less time will be required of the learner system and more time will be spent on the target system or process.

In general, future studies should focus on enhancing the relationship between the learning system and the target learning process in applications such as these, where the goal is to use the learning technology software system to learn a target software system. Many recommendations indicated that users did not want to spend an inordinate amount of time or effort on the learning system; they wanted to find the import information quickly and get back to their tasks. The learning system should be designed with that in mind; that it should work transparently for the users allowing them maximum time and effort for the target system or process. The target system should be the focus of the user’s cognitive load, not the learner system detailing how to use it.

Future studies should examine the effectiveness of iterative usability testing with other types of learning technologies and other content areas. The series of studies conducted can serve as a foundation for a formalized process/model for integration of iterative usability testing into the learning technology development process. Building such a model would be a great advance for the integration of usability testing into learning technology development process, but would in addition provide more general insights into the integration of usability testing in the design lifecycle more generally. In particular, queries into the following questions should be endeavored:

- What situations are most and least appropriate?
- What factors are most likely to lead to usability testing being effective or not effective as a part of the development process?

Related to this line of research, methods for most effectively “selling” usability testing as a viable component of the development process, and for educating team members in the role of usability testing as a component of development should be examined and identified.
3.5. CONCLUSION

This thesis explored the use of iterative usability testing applied to the design of a specific learning technology, a web-based learning system that will teach students to use Geographic Information Systems (GIS) within the foundational courses of a typical Civil Engineering program. Three separate instances of usability testing were iteratively during project development.

The first iteration involved the pre-design cycle, where existing software that dealt with similar design challenges was identified as Microsoft’s E-Learning Library (MELL). Participants in this study were required to create three increasingly difficult implementations of forms with MS Access. Four subjects who met the required content background were recruited for testing. The four subjects and their testing results indicated that the MELL design was an acceptable foundation on which to build the initial design. Problems for the system revolved around the video speed and navigation aspects. Recommendations were made to improve areas where MELL but also to draw out design aspects that MELL implemented well. The design team was given the recommendations and began production of the project prototype. Initial design choices were made and implemented with a portion of the content, leading to the second phase of usability testing.

The second iteration of usability testing followed the production of this initial prototype, and attempted to again ascertain strengths and weaknesses of the currently implemented design. Users with backgrounds similar to students who would use the final prototype were asked to make a map in ArcGIS with the aid of the learning system prototype. A number of interesting findings emerged from this testing. First, users had more trouble with the highest level scaffold of video due to its speed and the resulting amount of time required to watch it. Participants did not like waiting for video to complete tasks. This prototype was partial and did not include detailed text for task completion and the results from this testing on the video drew out the importance of clear and concise text. Linear navigation was found to be insufficiently implemented, with navigation controls outside of the users focus. Specific requests were made for a more salient linear style navigation to be implemented; since the linear navigation available went completely unnoticed by participants by and large. With these and other problems
in mind, recommendations for improved design were made and implemented in the next cycle of development.

The third iteration of usability testing came at the completion of this development, with a prototype that was nearing completion. This prototype was fully implemented and was tested as it will be in class/lab environment, with users working with the system to complete a geotechnical problem (borrow and fill site selection). This larger problem was given two-hours for completion, the same time as given to the deployment venue of the civil engineering lab. Five users who had taken the CE 215 class previously were recruited, with two users working as a pair (similar to lab option of partners). Overall, this system demonstrated the highest levels of usability within the iterations. User problems and comments focused on content versus system layout and design. Again video speed perturbed users for at times being too fast and others being too slow. The control provided for videos was insufficient. Recommendations were made for content corrections but for future installments to give the user stronger control for the video.

Taken together, the results from the testing iterations point to the importance of a learning tool as a flexible support system – the tool is not the task. This was evident even in the initial prototype study where learners were required to work with all components of the tool. Across all experiments learners tried to minimize their interaction with the tool by relying on minimal scaffolds as much as possible. Students were also most critical of, and performance was most hampered by, situations where they were unable to quickly locate the important information which would allow them to perform their primary task of carry out a task with the target software.

This leads to at least three general design principles, which apply in systems such as that developed in this project. First, the presence of multiple scaffolds is helpful since it allows the user to select the level of scaffold that is needed at any given time. Second, it is important that the learning system be flexible, including multiple types of navigation schemes (i.e., hierarchical and linear), and allowing for easy access of the necessary scaffold at any given time. Third, it is particularly important that the content be structured consistently and logically, allowing the learner to quickly develop an effective mental model, since the user needs to be able to quickly identify and acquire specific content.
As the iterations progressed, the usability and pedagogical effectiveness both appeared to improve, implying that iterative usability testing is a useful tool in the development of learning technologies. The growth of user acceptance through iterations is one indicator that usability testing integrated into the development process leads to more effective learning technologies.

However, much future work should be completed before this and other assertions can be strongly supported. The application of the prototype in its target class has occurred, and numerous data have been recorded. This data needs to be analyzed, interpreted and compared to the usability testing results, in order to provide additional insight into the efficacy of usability testing for evaluation of learning systems. More generally, researchers should look into learning issues such as the appropriate level of transparency of the learning system, the situational appropriateness of usability testing in learning technologies, and determining factors of usability testing effectiveness and their individual influences.
APPENDIX A: Subject Explanations for Questionnaire (Entrance & Exit)
Responses for MELL Existing System
This list contains each user’s explanation for how they rated a questionnaire inquiry (if provided) as well as their responses to the open-ended question (question 6) at the end of the exit questionnaire.

Participants’ quantitative answers to these questions are provided in Table 2.1.

Entrance Questionnaire:

I have a lot of experience working with databases
Subject A: IST 141 is the first time using
Subject B: Just began learning it
Subject C: I have worked with databases a little bit
Subject D: Just started Access in class

I have a lot of experience working with Microsoft Access
Subject A: IST 141
Subject B: Just began learning it
Subject C: All the work I have done with databases has been with Access
Subject D: Just started Access in class

I have a lot of experience working with Forms in Microsoft Access
Subject A: That is next week
Subject B: Just began learning it- sort of
Subject C: I have not used Forms
Subject D: Have not started forms yet

Exit Questionnaire:

I found it easy to complete the tasks from beginning to end
Subject A: Had trouble getting wizards to pop up when they should. No instructions for command buttons
Subject C: It was pretty easy. I had a little trouble getting the last one done, but not much
Subject D: The tasks were spelled out nicely

The MELL website helped me greatly in performing the tasks I completed.

Subject A: I know more about forms than without it
Subject C: I used it a little bit
Subject D: It had a friendly user interface

The MELL website was easy to use.

Subject A: Liked the layout and interface. Windows were big enough
Subject C: It seemed pretty simple from what I used

I relied a great deal on the website animations to complete the tasks

Subject A: to see where the cursor went since the graphics were not labeled
Subject C: I used the animation once to get me started, but it really didn’t show me what I wanted
Subject D: I used only one but it helped

I relied a great deal on the website text to complete the tasks

Subject A: Step instructions were useful
Subject C: I used the website text to give me a general direction to go in, and a layout of what needed to be done
Subject D: It was very informative.

What suggestions would you offer for improvement of the MELL interface so that it would make it easier for you to learn the materials?

Subject A: Higher contrast mouse cursor on example animations. Next step button slightly bigger and at top and bottom of window
Subject B: It’s difficult to understand, make it easier maybe…
Subject C: Possibly more descriptive text. I didn’t work with it enough to see if there were links to detailed descriptions for the tasks, but I think that would helped me, in an un-timed environment, get a firm grip on what it is I am actually doing.
Subject D: I would not do much different … I would try to make the tutorials open in the same window and not in a new one (Frames)
APPENDIX B: Subject Explanations for Questionnaire (Entrance & Exit)
Responses for Initial Prototype
This list contains each user’s explanation for how they rated a questionnaire inquiry (if provided) as well as their responses to the open-ended question at the end of the exit questionnaire.

Participants’ quantitative answers to these questions are provided in Table 2.2.

Entrance Questionnaire

I am familiar with ArcView GIS
Subject B: I’ve never used ArcView GIS.
Subject D: I have never used GIS

I am comfortable with ArcView GIS
Subject D: I have never used GIS

I am highly competent with ArcView GIS
Subject D: I have never used GIS

I am enthusiastic about GIS and Geotechnical engineering
Subject B: I do not know much about GIS, but I see it’s usefulness in real life settings
Subject D: I am interested in Geotechnical engineering

I have taken many soils classes
Subject B: I have only taken CE 215

Exit Questionnaire

I found it easy to complete the tasks from beginning to end.
Subject B: Some of the movies went too quickly; I had to go back a few times to catch up
Subject C: Maybe too easy – a little boring (move faster through video)
The provided website helped me greatly in performing the tasks I completed.

Subject A: I kept losing my place, but otherwise OK.

Subject B: Even though I had some trouble, I wouldn’t have gotten through the lesson without the website.

Subject C: At first, yes, but after I felt comfortable with the program I didn’t need as much details as the website gave

The website was easy to use.

Subject A: Initially I was a little lost on what I should do.

Subject B: It was pretty self-explanatory

I relied a great deal on the website animations to complete the tasks.

Subject A: The video(s) were very thorough

Subject B: The animations helped, but sometimes they moved too fast or too slow.

Subject C: At first, but not towards the end

I relied a great deal on the website text to complete the tasks.

Subject B: The bulleted lists were better than the paragraph-form text. It more or less just helped explain what was going on.

Subject C: Video was more helpful

What suggestions would you offer for improvement of the website interface so that it would make it easier for you to learn the materials?

Subject A: Wanted some way to show one’s current location on the sidebar, or “some way of keeping track of what you should click next.” Requested bigger text.

Subject B: If the website could suggest to the user to adjust the size of the windows so the user could see both the website/movie and the GIS. Possibly make the website movies continue on after the user clicks a button on the movie (so it knows it can proceed).
Subject C: Better Introduction. Clickable (skip-able) control of video
Subject D: I found the site very helpful and I personally can’t think of anything that would make it easier
APPENDIX C: Subject Explanations for Questionnaire (Entrance & Exit)

Responses for Final Prototype
This list contains each user’s explanation for how they rated a questionnaire inquiry (if provided) as well as their responses to the open-ended questions (13 & 14) at the end of the exit questionnaire.

Participants’ quantitative answers to these questions are provided in Table 2.3.

Entrance Questionnaire:

**I am familiar with ArcView GIS**

Subject A: I might have used this before.
Subject B: Never worked with it
Subject C1: Not sure what it is
Subject C2: I’ve never heard of it

**I am comfortable with ArcView GIS**

Subject B: Never worked with it
Subject C1: Never used it
Subject C2: I haven’t ever used it.

**I am highly competent with ArcView GIS**

Subject B: Never worked with it
Subject C1: Never used it
Subject C2: Again, I haven’t used it.

**I am enthusiastic about GIS and Geotechnical engineering**

Subject B: I really like soils and programs that go along with them
Subject C2: I’m not sure what GIS is.

**I have taken many soils classes**

Subject B: CE 315, CE215, CE229
Subject C1: Presently taking foundations
Subject C2: I am in the process of taking my 2nd soils class.
Exit Questionnaire:

I learned a great deal of information about soil borrow site selection from this week’s lab.

Subject A: You really get to see what you would have to work with in industry
Subject B: I think it’s a good tool to use.
Subject C1: I thought the program did a good job about borrow site selection.
Subject C2: Worked out well by showing borrow sites in relation to the construction site.

I found this week’s lab on soil borrow site selection to be very motivational.

Subject B: I thought it was neat to use a program like this.
Subject C2: It had real world application

This week’s lab activity over soil borrow sites was applicable to “real world” engineering.

Subject B: I would think that engineers go through steps like this when calculating borrow and fill quantities and prices.
Subject C2: It was applicable.

Before the lab activity that covered soil borrow sites, I knew a great deal about the subject.

Subject B: Not really much
Subject C2: Learned a lot from already having taken the course.
Subject D: Family is in excavating business, had some prior experience.

After the lab activity that covered soil borrow sites, I knew a great deal about the subject.

Subject B: I learned a little bit more about it.
Subject C2: It made the concepts easy to grasp.
Please list the strengths of the lab activity that covered soil borrow sites, in terms of it’s effect on learning and motivation, and it’s applicability to “real world” engineering.

Subject A: I thought that the program was great because it had all the tools that you need. That program was really applicable.

Subject B: It seems like a really useful and informative program. I would actually like to see more of these programs taught and used by students.

Subject C1: Shows what factors are taken into account when determining what soil borrow site to choose.

Subject C2: It was enjoyable and easy to follow.

Subject D: Actual maps and sites gave more of a real-world feel than “box & line” diagrams used previously in 215. Lab required use of all major real world variables involved in borrow site selection (material cost, transportation cost, available materials).

Please list ways in which the lab activity that covered soil borrow sites could be improved.

Subject A: Information on which type of soil you want for the foundation and unit weight for all [could be provided]

Subject B: Make the instruction[s] a little more clear in certain areas.

Subject C1: I think the units where it said the truck carried 15 tons was supposed to be 15 CY, because we needed 880 CY and didn’t know how much CY [cubic yards] a truck carried. The very last section on “Proposed problem” seemed like it could have been up further in the outline, because we needed that to make our choice on what borrow site to choose.

Subject C2: Some of the steps seemed redundant. It wasn’t always easy to find the information required to solve the problems. Maybe get rid of some of the extra layers.

Subject D: None come to mind.
APPENDIX D: Technically Oriented Questions Administered Concurrently with the Final Prototype Exit Questionnaire
This page details the technical questions asked of the production prototype study participants. These questions were administered mainly for timing issues, to see if the questionnaire packet for the full experiment, which was following soon after, was a reasonable size.

Participants overall score for these questions is located in Table 3.

**Technical Oriented Questionnaire (have we learned something?):**

1. A borrow site is always located at quarries (T/F)
2. A rock quarry could serve as a borrow site if granular fills are desired. (T/F)
3. The acronym GIS stands for: Geologic Inspection Standards. (T/F)
4. The following disciplines make use of GIS:
   a. City Planning
   b. Water Resources
   c. Geology
   d. Anthropology
   e. All of the above
5. Which of the following is not needed to estimate the cost of imported soils to a site:
   a. Delivery cost
   b. Cost of material per cubic yard
   c. Soil type
   d. Compaction testing
6. The geology at a site is not important when making a selection for soil borrow sites. (T/F)
7. GIS can be used for the following:
   a. Composing letters
   b. Purchases online
   c. Locating sites
   d. Soil Testing
8. Results of the Plastic and Liquid Limits can be obtained without running lab tests. (T/F)
9. If fill is required for a construction site, the soil type is not important as long as there is enough material available at reasonable cost. (T/F)
10. The Plastic and Liquid limits are important geotechnical lab tests to run on a granular backfill. (T/F)
11. The usefulness of GIS in geotechnical projects lies in the spatial analysis and attribute storage capabilities of the GIS. (T/F)
12. Other factors that may increase costs when a material is used at a project site are:
a. Labor Costs
b. Equipment Costs
c. Shrink / Swell of the material
BIBLIOGRAPHY


Hall, R. H. (2002). An integrative approach for teaching web design. Minneapolis, MN.


American Society for Engineering Education:


VITA

John McKenna Sullivan was born in St. Louis, Missouri, on March 26, 1981. In May 2003, he received his B.S. in Management Systems from the University of Missouri – Rolla with minors in Computer Science, Psychology, and Economics. He has worked as a graduate research assistant from September 2003 till December 2004 in the department of Information Science and Technology (IST). A large portion of this tenure was spent in the development of the GIS Learning Object project during which his first journal paper was published. In December 2004, he received his M.S. in Information Science & Technology with honors from the University of Missouri – Rolla.