2007

Proximity Navigation for Map-Based Interfaces: Generalizing Menu Design for Multiple Dimensions

Brian Scott Malek
University of North Florida

PROXIMITY NAVIGATION FOR MAP-BASED INTERFACES: GENERALIZING MENU DESIGN FOR MULTIPLE DIMENSIONS

by

Brian Scott Malek

A thesis submitted to the
School of Computing
in partial fulfillment of the requirements for the degree of

Master of Science in Computer and Information Science

UNIVERSITY OF NORTH FLORIDA
SCHOOL OF COMPUTING

April, 2007

The thesis "Proximity Navigation for Map-Based Interfaces: Generalizing Menu Design for Multiple Dimensions" submitted by Brian Scott Malek in partial fulfillment of the requirements for the degree of Master of Science in Computer and Information Sciences has been

Approved by the thesis committee: Date

Signature Deleted

______________________ 5/3/07

F. Layne Wallace
Thesis Advisor and Committee Chairperson

Signature Deleted

______________________ 5/3/07

Sherif A. Elfayoumy

Signature Deleted

______________________ 5/3/07

Arturo J. Sanchez

Accepted for the School of Computing:

Signature Deleted

______________________ 5/8/07

Judith L. Solano
Director of the School

Accepted for the College of Computing, Engineering, and Construction:

Signature Deleted

______________________ 5/10/07

Neal S. Coulter
Dean of the College

Accepted for the University:

Signature Deleted

______________________ 10 MAY 2007

David E. W. Fenner
Dean of the Graduate School

ACKNOWLEDGEMENT

I would like to thank four people, without whom this would never have been accomplished. First, I would like to thank Bernard Sanders, Jr. for allowing a flexible work schedule that enabled me to conduct meetings at school during normal working hours. Second, I would like to thank Wesley Page for tirelessly testing the implemented systems and searching for flaws. Third, I would like to thank my thesis advisor, Dr. F. Layne Wallace, for his constant support of the work done throughout this process. Finally, I wish to especially thank my wife Jamie Malek for her support and understanding during the many hours spent apart to achieve this milestone in my life and academic career.

CONTENTS

FIGURES

TABLES

ABSTRACT

The development of effective multidimensional map-based interfaces is an important area of research in need of design techniques and guidelines. To date, guidelines for multidimensional interfaces have been generalized from text-based interfaces and few experimental studies have been conducted to asses their effectiveness.

Guidelines for design were studied with the goal of extending the current body of knowledge about the usability of these interfaces. Based on design guidelines, multidimensional map-based interfaces with various levels of depth and breath, with and without scent-based components were used to perform simple and compound tasks. The goal of this study was to investigate the effectiveness of design guidelines on response time, preferences, and navigation and task accuracy.

Results showed relationships exist among navigation and task accuracy, response time, and preferences within simple or compound tasks. However, few relationships exist between simple and compound tasks. Contrary to results from previous research, interface depth and breadth was found to have no significant effect on navigation and task accuracy or response time. For compound tasks, interfaces with scent-based components were found to be more effective regarding task accuracy at greater depth levels. The absence of scent in the interface was shown to be more efficient regarding response time and navigation accuracy during compound tasks.

Chapter 1

INTRODUCTION

While many have studied the use of text-based hierarchical menu systems, little has been done with multidimensional graphical interfaces. One type of multidimensional graphical menu system that has been studied, and continues to be an important area of research, is the map-based interface [Hornbaek02]. Creation of multidimensional map-based interfaces is made attractive to developers by the increasing amount of geospatial information on the World Wide Web [Lim02]. However, while research has been performed regarding zooming of the interface, few empirical studies have been found dealing with the issues of depth versus breadth, scent, or task complexity.

A common technique used in human-computer dialogues is the hierarchical menu structure [Gray86], which contains a series of menus with a main menu and a number of submenus. Early human-computer dialogues were text-based in nature and limited in display size, therefore the use of a hierarchical menu system served two purposes: information could be broken up onto several screens to save screen space, and information could be categorized into organized units. Many studies have been performed [Gray86, Larson98, Tullis85] regarding the issue of depth versus breadth in hierarchical menu systems, yet the results are as varied as the studies done. With the introduction of the graphical user interface (GUI) and the World Wide Web, one might believe the hierarchical menu system would not survive, but many popular Web-based systems include hierarchical menu systems.

CNN's website (http://www.cnn.com) is a good example of a hierarchical menu system. The front page contains a list of submenus (Home, World, U.S., Weather, Business, Sports, Analysis, Politics, Law, Tech, Science, Health, Entertainment, Offbeat, Travel, Education, Specials, Autos, I-Reports) from which a visitor can choose. Each of these submenus displays further submenus when chosen. For example, the selection of Sports not only brings visitors to a main page for sports news, but also displays submenus for specific sports. Other examples include customizable search engine interfaces such as Google (http://www.google.com) and Yahoo! (http://www.yahoo.com), which allow users to place chunks of information on a single page which leads to more information about each topic.

Another aspect of hierarchical menu systems which has been studied is the role of menu titles as a navigational aid [Gray86]. Menu titles are used in hierarchical menu systems to aid a user in menu navigation. They normally provide the user with information regarding the menu they selected as well as information regarding the menu previously displayed. The use of menu titles is similar to the concept of 'scent' or the information a user can derive from a structure's design and the relative location of the target [Larson98].

1.1 Multidimensional Map-Based Interfaces

According to Lim [Lim02], the increasing amount of geospatial information on the World Wide Web makes it attractive to developers to create portals that organize these resources spatially on a map via a graphical user interface. Lim developed a multidimensional map-based user interface, named G-Portal which allowed users to visualize distributed data in the context of a map where users could locate information based on location. The aim of G-Portal was "to identify, classify, and organize geospatial and georeferenced resources on the web and to provide digital library services (e.g. searching and visualization) for these types of resources," however no empirical data was gathered to measure the effectiveness of Lim's interface.

The main method for accessing information on G-Portal was through a multidimensional map-based interface. Users navigated through the interface by using a set of navigation tools – including zooming. Information displayed on the interface could be turned on or off using a series of layers. All information displayed on the map had a corresponding layer. A series of checkboxes were used to allow users to display the information on the map by selecting, or deselecting, the appropriate checkbox. According to Lim [Lim02], G-Portal is just one of many similar interfaces whose goal is to provide information based on geospatial and georeferenced context. Other systems include Georep and the Spatial Document Locator System, which provide search services for geospatial data on the Web.

1.2 Zoomable User Interfaces

According to Hornbaek [Hornbaek02], the creation of systems for information visualization has become a successful methodology for human-computer interaction. However, few empirical studies have investigated the usability of zoomable user interfaces. Many of the systems previously defined contained an overview of the interface, or a separate view with a zoomed-out image of the current view. In his study, Hornbaek designed an experiment to investigate the impact of zoomable user interfaces with and without an overview on usability and navigation patterns. However, Hornbaek did not examine the impact of menu titles or labels in his experiment.

Hornbaek considered two characteristics of a zoomable user interface: objects are organized with reference to space and scale and users interacted directly with the information by panning or zooming. Hornbaek also introduced the concept of semantic zooming, where areas on the map could be shown with different features or details, such as county names, cities, or borders, depending on the scale [Hornbaek02].

Several methods of zooming were also defined, including goal-directed zooming, the combination of zooming and panning, and automatic zooming. In goal-directed zooming, zooming occurs to a specific scale. In combination zooming and panning, extensive panning from side to side leads to zooming. And finally, in automatic zooming a click of the mouse determines the area zoomed in upon. Two ways of implementing zooming in the interface were also introduced: jump zooming and animated zooming. Jump zooming occurs instantaneously from one scale to the next, whereas in animated zooming the change occurs smoothly over a set period of time. Hornbaek referenced a previous study done on the two zooming techniques and noted subjects performed better at reconstruction of the navigation using animated zooming, but no difference in satisfaction or time was found [Hornbaek02]. Results vary based on the different types of interfaces used since each experiment varied in how zooming was implemented as well as the amount of information used.

In Hornbaek's experiment, subjects used map-based interfaces to solve tasks on two differently organized maps. The author cited three reasons for using maps for the experiment: map interfaces constitute an important area of research, maps include characteristics of other commonly used information structures, and the direct relationship between representation and physical reality aids in interpretation [Hornbaek02]. The experiment consisted of 10 tasks, five based on navigation and five based on browsing. Navigation tasks required subjects to find an object, or multiple objects, on the map. Two navigation tasks required finding a single object, two required finding multiple objects, and one required finding the route between two objects. Browsing tasks required the subjects to scan the entire map for certain types of objects. During the experiment information was gathered regarding the accuracy of the questions asked, task completion time, preference, satisfaction, and navigation actions. The results of the experiment showed a direct manipulation interface can reduce or even eliminate the need for a separate overview [Hornbaek02].

1.3 View Navigation

Effective view navigation was introduced by Furnas, who stated that information which helps determine where to go next is central to view navigation [Furnas97]. The author argued "despite the vastness of an information structure, the views must be small, moving around must not take too many steps and the route to any target must be discoverable." The beginning of the Web found a richness of information available to users unlike anything else before, yet navigation from one place to the next was a series of sometimes unrelated hyperlinks leading the user from page to page. The difficulty of information gathering on the Web led to the rise of search engines and pure navigation is now a thing of the past.

While Furnas' work was mainly theoretical, it provided some insight into issues such as navigation and finding data in information structures, where the focus was on very large systems when time and physical resources are limited. Two terms were introduced: view traversal and view navigation. View traversal is the iterative process of viewing, selecting something, and moving to it to form a path. View navigation attempts to ease the traversal path to a target by providing reasonable and informed information on the path and target [Furnas97]. The author argues structures like the World Wide Web are bad examples of view navigation, while semantic zooming is a good example because it provides better information with regards to the desired target.

1.4 Hierarchical Menu Systems

Menu systems have always aimed at being as user friendly as possible with respect to the ability to find information. One of the most common menu systems is a hierarchical system which consists of a main menu and a series of submenus [Gray86]. A commonly studied aspect of hierarchical menu systems is the issue of depth versus breadth, where depth is defined as the number of submenus available and breadth is the number of available options to choose from and their impact on information retrieval. Various studies have been performed and many of the results contradict one another. Research found on hierarchical menu systems focused on text based links. The World Wide Web contains many types of hierarchical menu systems, many of which are based on images, or text inside images, yet no research was found on the issue of depth versus breadth in this type of environment.

In an experiment performed by Tullis, the author aimed to answer the question of depth versus breadth regarding logically related information [Tullis85]. Based on results from previous research, an experiment was designed where logical relationships were determined among a series of commands and two separate hierarchical menu systems were created based on varying depths and breadths. The first system was a narrower and deeper single-column menu and the second was a wider and shallower multi-column menu. The multi-column menu would allow up to three columns and a maximum of 45 selections per menu where the single column menu was limited to a maximum of 15 selections.

The experiment consisted of a series of 24 tasks to accomplish using one of the two interfaces – a subject would not use both interfaces. Subjects were asked to perform the tasks and record any output displayed. Three metrics were recorded for each subject: total number of steps taken, total time to accomplish the set of tasks, and the total number of errors committed. The results of the experiment showed significantly more extra steps were taken using the single-column menus than the multi-column menus. However, despite the extra steps, subjects did not take significantly longer to accomplish the tasks. Since subjects performed the same series of tasks in essentially the same time and subjects were able to predict which path to take in broader structures, Tullis concluded designers should strive for breadth over depth when creating hierarchical menu systems [Tullis85].

Landauer performed a similar experiment with regard to depth and breadth, but aimed to answer whether menu choice response time is determined by a choice among responses, or the visual scan-and-match process [Landauer85]. In the experiment, subjects were required to select one word or number by successive choices among ranges of words or numbers. Words, which consisted of words found in a dictionary which were four to 14 characters long, were displayed in alphabetic order and numbers, which were the natural numbers one to 4096, were displayed in numerical order. For each of the questions the word or number to be found would be displayed on an otherwise blank screen, then when the subject was ready it would be displayed on a second screen located above and to the left of the menu system which the subject was using to find it. Words and numbers were selected so they would never appear until the last screen, or lowest depth level.

Subjects were separated into two sets: those who searched for numbers then words and those who searched for words then numbers. To equalize the number of selections for each of the depth levels, a predetermined number of trials were executed with a set number of selections per trial. The results showed the penalty for a deep menu was large with respect to time where in some cases the overall time to find the target in a deep menu was twice as long as a broad menu. According to Landauer, the results showed "it is clear that in the choice situation studied here broader, shallower menu trees yield a faster search than narrow, deeper ones" [Landauer85]. However, the author noted the categories themselves may have some impact on the results. For example, "consider dividing the United States into 3 or 50 versus 25 geographical categories," [Landauer85] the results may vary based on how the categories are divided. The author also stated more results from similar experiments would be needed before a generalization could be made regarding design methods.

1.4.1 Scent for Information Retrieval in Hierarchical Menu Systems

Both Landauer [Landauer85] and Furnas [Furnas97] concluded the organization of the desired target information may have an impact on information retrieval. Larson discusses a similar concept known as 'scent' or conveying target information via category labeling. Scent in a menu structure can be made via category and subcategory labeling. In this way if the category of 'Science News' is selected from an upper-level menu, it would have several subcategories such as 'Science News: Physics' or 'Science News: Psychology' [Larson98]. The goal is to convey as much information as possible about the structure's design and the location of the target via labeling. Larson designed an experiment using the notion of scent by designing a categorization scheme from an encyclopedia with varying depths and breadths. While Larson categorized items into logically organized nodes, no research was found determining the impact of scent on physically organized interfaces such as a map.

Larson's experiment consisted of three separate hierarchical menu systems of hyperlinks with a total of 512 nodes. However, unlike other experiments where organization was done from the bottom up, the three structures were created with items naturally belonging to those structures. Due to this restriction on design, only 128 of the 512 items appeared on all three structures. In the experiment, only these items were set as possible targets. To ensure optimal scent, the top page of each hierarchy was labeled using the same naming scheme: 'hierarchy 1:,' 'hierarchy 2:,' or 'hierarchy 3:' and each of the second level pages were named 'hierarchy x: appropriate page title:' [Larson98]. Under the category name was a vertical list of items in random order which were displayed in either single-column or multi-column format.

Subjects were given 24 total searches to perform, eight in each structure. During the experiment three metrics were collected: accuracy, completion time, and preference. After the subjects completed the 24 searches, a five point Likert scale was used to determine preference based on several questions regarding the hierarchies. Reaction time and accuracy metrics showed a broader, shallower hierarchy performed the best. However, out of the two broad, shallow hierarchies used, the one which had the best accuracy and time metrics was not the one the subjects preferred. Larson concluded "our findings are consistent with those reviewed earlier that favored breadth over depth, even with our structures that were expertly organized to deliver optimal scent" [Larson98].

1.4.2 The Role of Menu Titles in Navigation

Related to the role of scent, the role of menu titles can aid in orienting subjects during navigation [Gray86]. Gray designed an experiment to determine the effectiveness of the presence or absence of menu titles on search time and error rate in hierarchical menu structures. Search time was measured from the initial display of the target item until it was selected from the list. Gray used a similar naming scheme to Larson, where if 'Animal' was selected from the main menu, the submenu would read 'Main Menu – Animal' [Gray86].

Gray noted there was no statistically significant difference between the total times of the two systems or in the number of errors committed. However, once depth reached levels greater than three, the group with menu titles made fewer incorrect choices. Gray concluded the presence of menu titles did not affect time or error rate, yet it did have an impact at lower depth levels where subjects made 19% more errors at these levels than those who had menu titles [Gray86]. While Gray's experiment measured the effect of menu labels on a text-based hierarchical menu system, research has not been conducted to study the effect of labels in a map-based interface.

While systems were developed as multidimensional graphical map-based interfaces, little empirical data was collected to determine their effectiveness. Much of the research done on hierarchical menu systems regarding depth versus breadth, scent, and menu titles were based on systems with a series of logically organized nodes not physically organized constructs. In the current research, issues of scent (menu titles), depth versus breadth, and task complexity in a multidimensional graphical menu system were examined.

Chapter 2

METHODOLOGY

There are many characteristics and facets regarding multidimensional graphical user interfaces. A geographical map was chosen as the construct for this experiment. Automatic zooming was used for navigation and a study was done to measure the effects of depth versus breadth, scent, and menu titles on response time, accuracy, and preference. Depth is defined as the number of times the interface can be zoomed in and breadth as the number of proximal areas to choose at a given depth level. Optimal scent was achieved by the use of menu titles, where information about the previous depth level could be displayed, as well as by the use of labels on the map itself. Based on the prior research, a series of multidimensional graphical map-based menu systems were designed and built to measure subject response time, accuracy, and preferences.

2.1 Menu Design of a Multidimensional Map-Based Interface

The first step in the design of the graphical multidimensional map-based menu system was to locate a map of the world which could be resized and still retain image quality in a Web-based browsing environment. The Central Intelligence Agency's World Factbook [CIA06] contained two maps of the world in vector graphic format. The first was the physical map of the world which was ruled out because it did not show enough cities and contained topographical information which was considered distracting. The second was the political map of the world which contained more cities and was designed to have a more two-dimensional look which provided less distraction. Macromedia Dreamweaver 8 [MacromediaDreamweaver05] was used to create two distinct types of menu systems: one with map labels and menu titles where optimal scent would be provided to subjects and one without labels or titles. Figure 1 shows the upper-most level of the label based menu while Figure 2 shows the upper-most level of the non-label based menu. Three distinct subsystems based on varying depths and breadths described below were created from these two designs.

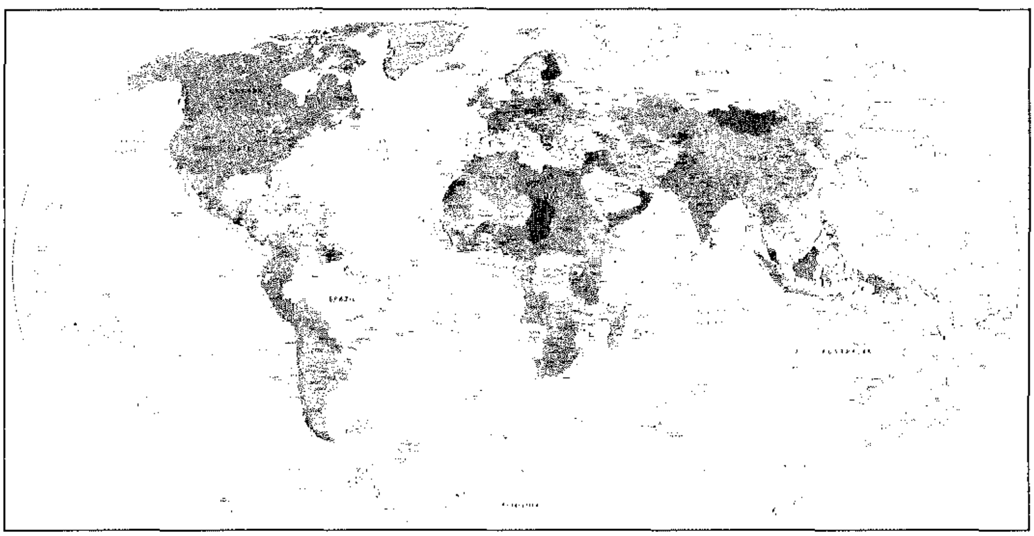

Figure 1: Uppermost Level of Label Based Graphical Menu

The size of the interface was determined by the default resolution on the LCD monitors – 1280x1024 pixels. Due to this constraint the maximum width of the map was determined to be 1000 pixels. The aspect ratio of the map was determined to be 25 pixels wide to 13 pixels high. Using the aspect ratio of the map, the maximum height of the map portion of the interface was calculated to be 520 pixels. The remaining pixels were taken up by the browser interface and also used to display the question and possible answers for the subjects. Since the interface was designed to be accessed via a Web browser it was determined that no scrolling should take place.

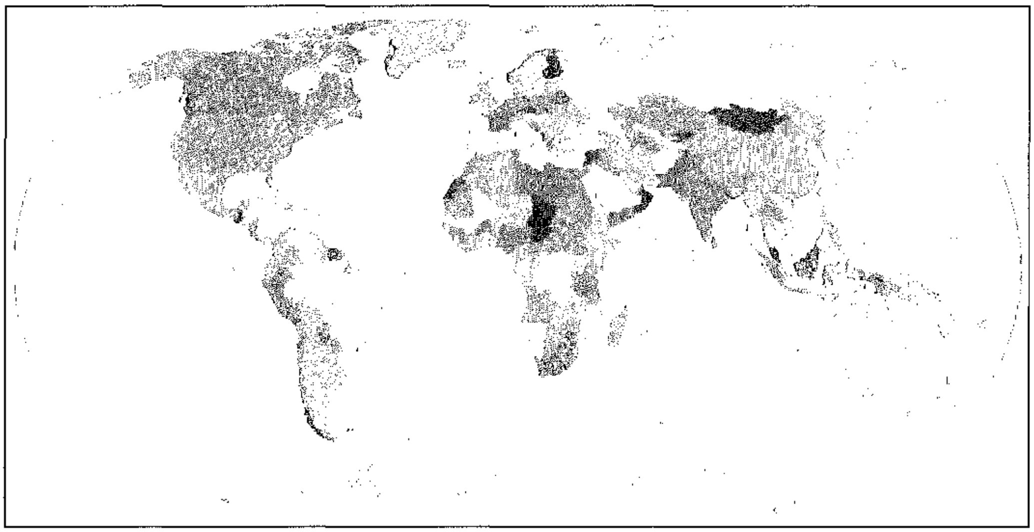

Figure 2: Uppermost Level of Non-Label Based Graphical Menu

2.1.1 Depth and Breadth

Miller noted "by organizing the stimulus input simultaneously into several dimensions and successively into a sequence or chunks, we manage to break (or at least stretch) this informational bottleneck" [Miller56]. In other words, it is possible to increase the amount or size of information one can store in short-term memory by organizing the information into a sequence or a logically organized group. Shneiderman [Shneiderman05] best explains this concept:

> "Most Americans can also probably remember seven decimal digits, seven alphabetical characters, seven English words, or even seven familiar advertising slogans. Although these items have increasing complexity, they are still treated as single chunks. However, Americans might not succeed in remembering seven Russian letters, Chinese pictograms, or Polish sayings. Knowledge and experience govern the size of a chunk and the ease of remembering for each individual."

This limitation in human short-term memory was considered during creation of the menu system by examining depths and breadths of the menu structures in prior research.

Depths and breadths of the menu system were chosen based on previous experiments in hierarchical menu structures. Gray studied a 4x4x4 (4^3) interface [Gray86], Landauer studied breadths of 2, 4, 8, and 16 [Landauer85], while Larson studied an 8x8x8 (8^3) interface [Larson98]. To ensure consistency between interfaces, the total number of possible choices (depth * breadth) had to remain consistent [Larson98]. Kiger [c.f. Larson98] studied five menu structures: 2^6, 4^3, 8^2, 16x4, and 4x16 while Zaphiris and Mtei [c.f. Larson98] replicated Kiger's structures using web hyperlinks. Due to the constraint of total number of choices remaining consistent, three separate breadths and two separate depths were selected: 4^3, 8^2, and 16x4. Figures 3, 4, and 5 show how the map of the world, or main menu of the interface, was divided into regions for the varying breadths. Both the label and non-label based menu systems were divided in the same manner. For the 4^3 and 8^2 interfaces the division of regions remained consistent at the lower depth levels. However, at the second depth level for the 16x4 interface, the resulting region was divided in the same way as the 4^3 interface.

The division of the map into regions presented a problem for the 8^2 interface. Both four and 16 are perfect squares so when the region is divided the aspect ratio of the image remains the same. Eight is not a perfect square; therefore the subsequent divided regions did not contain the same height to width aspect ratio as the other two interfaces. Because the resulting width to height aspect ratio was 50 to 13, the menu system for the 8^2 interface was designed with images which were 1000x260 pixels in dimension. The 4^3 and 16x4 interfaces were designed with images which were 1000x520 in dimension.

Since subjects would only interact with one interface, it was determined the difference between the aspect ratios would have no effect on the interaction of the subject with the interface.

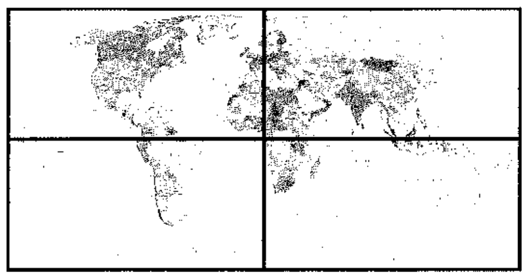

Figure 3: Regions for the 4x4x4 Interface

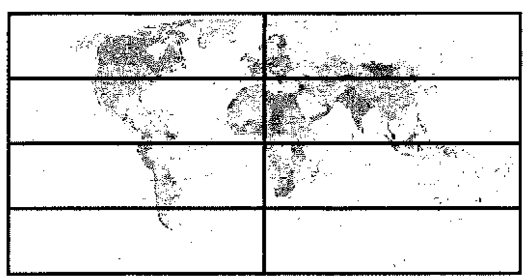

Figure 4: Regions for the 8x8 Interface

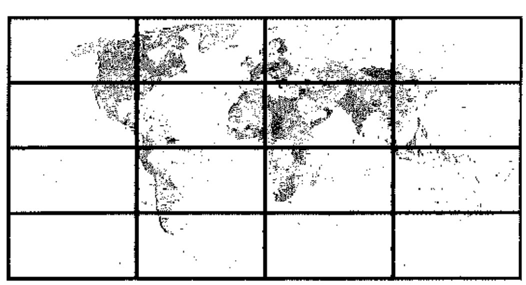

Figure 5: Regions for the 16x4 Interface

2.1.2 Creating the Label Based Regions

Since the original image was in vector format, Adobe Illustrator CS2 [AdobeIllustrator05] was used to open and resize the image. The image was resized in scale to 8000x4160 pixels, which was large enough to create the 64 total regions for each of the interfaces. The image was then imported into Adobe Photoshop CS2 [AdobePhotoshop05] where guides were used to divide the image. For the 4^3 interface a vertical guide was created at 50% height and a horizontal guide was created at 50% width creating a 2x2 section. The individual sections were then copied into new images to be further divided into another 2x2 section, which was further divided into a final 2x2 section. Each of the individual sections was copied into their own image and saved using a consistent naming scheme. The top-left region was named R1, the top-right R2, the bottom-left R3, and the bottom-right R4. When R1 was subdivided the resulting regions were named R1_R1, R1_R2, R1_R3, and R1_R4. When R1_R1 was divided the resulting regions were named R1_R1_R1, R1_R1_R2, R1_R1_R3, and R1_R1_R4. The same methodology was repeated to make the regions for the 8^2 and 16x4 interfaces.

2.1.3 Creating the Non-Label Based Regions

Creation of the non-label based regions followed a similar pattern. Within Adobe Illustrator [AdobeIllustrator05] the labels for each of the cities, countries, oceans, and other extraneous information were first removed from the map. The map was then imported into Adobe Photoshop [AdobePhotoshop05] where the same process used to create the label based regions was used to create the non-label based regions. The same naming scheme was also used. At the lowest depth level the label based images were used as this allowed subjects to find the city in question. Since both images were created from the same original image, the lowest deptb levels matched exactly to the label based interface.

2.1.4 Interface Design

Once all of the images were created, the interface was created so interaction with the interface could be tracked. For tracking purposes an SQL-based database (MySQL) was created and web-embedded script languages (PHP, JavaScript) were used to connect to the database and dynamically generate the pages. The final task was to retrieve information about cities around the world. This information was used as responses to questions during the study. Once the information was gathered, the pages were built using Macromedia Dreamweaver 8 [MacromediaDreamweaver05].

2.2 Tasks and Metrics

Before interface design could proceed the tasks were identified as well as a means to measure each task. Using the map-based interface subjects were given two distinct types of tasks to perform: simple and compound. Simple tasks consisted of using the interface to find information about a particular city in the world. Compound tasks consisted of using the interface to compare information gathered about two cities in the world. To remain consistent and to avoid confusion regarding the goal of the question, each of the compound tasks was phrased in the same fashion so subjects would always be searching for the larger of two numbers. Out of the 10 total questions for which the subjects must find an answer, five were simple and five were compound. Appendix A contains a list of the simple questions and Appendix B contains a list of the compound questions.

Three different metrics were collected in the study: response time, accuracy, and preferences. Response time contained two separate metrics: the time it took to make the first click, and the overall time spent on a question. Accuracy was measured in two different ways: total number of navigation errors committed and whether or not the subject answered the question correctly. Preferences were measured on a five-point Likert scale and were asked at the completion of the 10 questions.

2.2.1 Response Time

To calculate response time, two separate times were collected: initial choice time and total time spent on each question. The dynamically generated pages were designed to aid collection of these two times by the use of Web-embedded script languages. The first screen of the interface gave information regarding the study and contained a 'Begin' button for when the subjects were ready to start. When the button was pressed, the subjects were presented with the question as well as another 'Begin' button. When this button was clicked, the main menu was displayed. The main menu contained the question to be answered at the top along with five possible answers as radio buttons if the question was simple. If the question was compound, two radio buttons were present. The main menu also contained a 'Continue' button to confirm the subject's choice and the initial map of the world. Figure 6 shows an example of the question on an otherwise blank screen and Figure 7 shows the same question with the possible answers and the initial map of the world. Figure 8 shows the initial map of the world for compound questions.

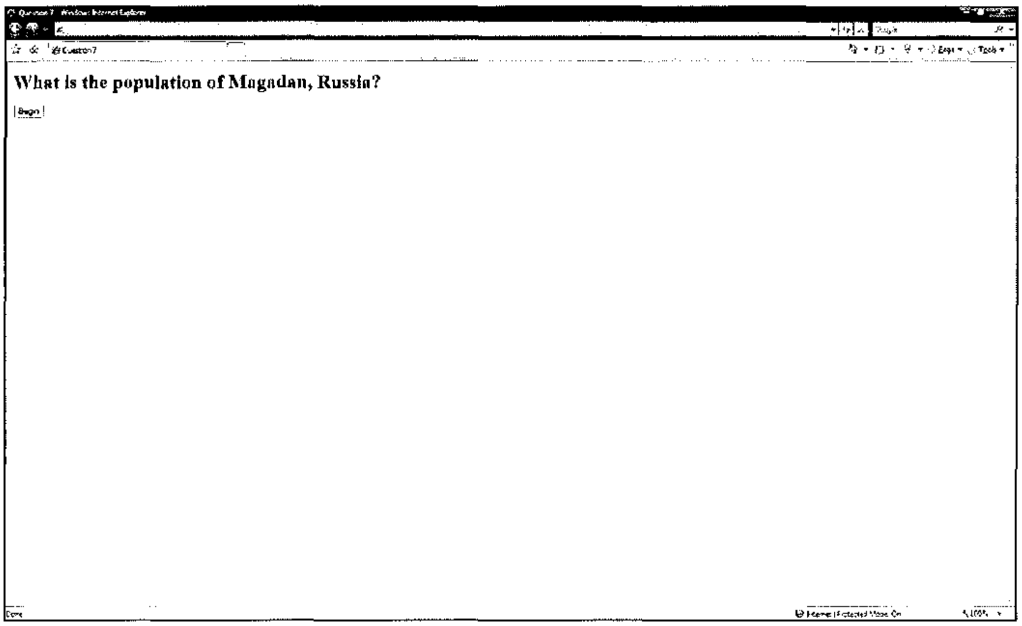

Figure 6: Question and Begin Button

When the 'Begin' button was selected on the question page, the timestamp was saved to the database. The timestamp when the subject made their initial choice by clicking on the map interface was also saved. The difference between these two times was measured as the initial choice time. After the subject found the target city, or cities, an answer was selected from the radio list and the 'Continue' button was selected to continue to the next question. The timestamp for this transaction was also saved and the subject was directed to the next question. The difference in the time when the 'Continue' was selected to answer the question and when the 'Begin' button had been selected at the beginning of the question was measured as the time spent on the question. Subsequent questions repeated the same process.

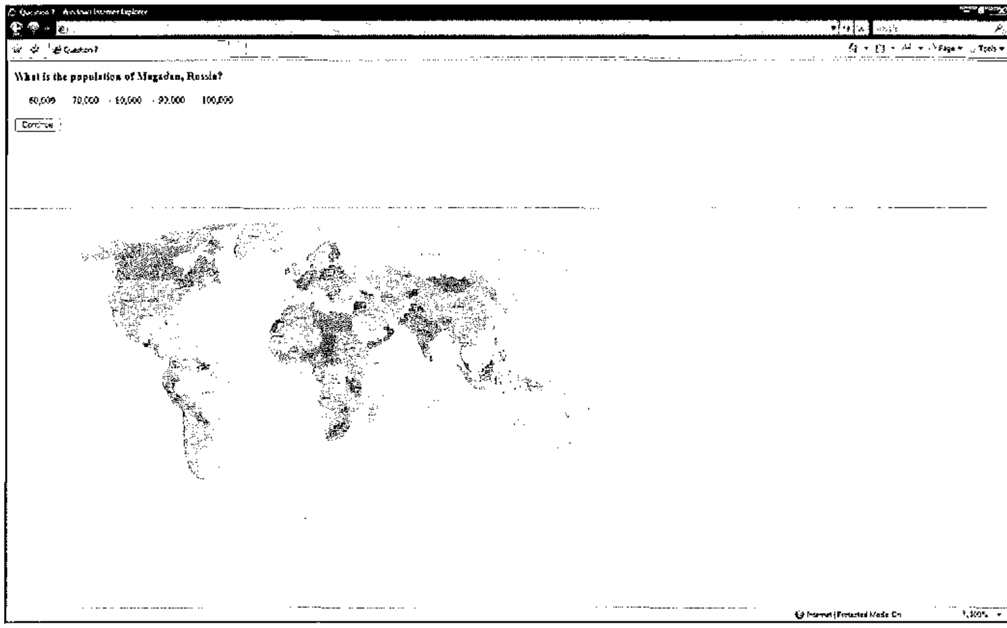

Figure 7: Simple Question, Possible Answers, and Initial Map

Figure 8: Compound Question, Possible Answers, and Initial Map

2.2.2 Accuracy

Accuracy was another non-subjective measurement collected. Similarly to response time, accuracy had separate components. The first component was the number of navigation errors made using the map interface and the second was whether or not the subject answered the question correctly. The number of navigation errors was determined by the number of selections a subject made which were not along the correct path. If the subject chose the wrong region, it was counted as one navigation error. However, if the subject zoomed out because they realized they made the wrong choice, it was not considered an error as this was the correct action to take to get to the target. For example, assume the correct path for a subject to take in the 4x4x4 interface from the main menu was: R1, R1_R2, and then R1_R2_R4. The subject took the path: R2, Main Menu, R1, R1_R2, and then R1_R2_R4. The subject made one error, the initial choice of R2 instead of R1. The subject choosing to go back to the Main Menu was considered correct as it is the only way to get back to R1.

The final accnracy measurement was whether or not the subject answered the question correctly. It was not the intention of the questions to be difficult to answer, however, task accuracy was used to examine relationships among preferences, navigation accuracy, and overall response time.

2.2.3 Preference

The final measurement collected was subject preference. At the completion of the 10 questions the subject was given a series of demographic questions as well as a set of preference questions based on a five-point Likert scale with 1 being 'Strongly Disagree' and 5 being 'Strongly Agree.' Subjects were asked questions regarding the ability to find information using the interface, the organization of the interface, their confidence using the interface, and their level of liking of the interface. During the collection of preference questions demographic questions were also asked. These questions included computer experience, handedness, eyesight, age, and gender. For a full list of preference and demographic questions see Appendix C.

2.3 Selecting the Cities

Cities were selected by examining each of the regions created during menu design. Since the object of the study was not to test the subject's knowledge of geography, many of the cities selected were country capitals or cities in the news. For each of the cities two items were retrieved: the population, and a random fact about the city. The random fact could be the date it was established, the length of the river it lies upon, its population density, or any number of other facts containing a number. Facts based on numbers were selected because subjects were asked to compare information about two cities and the comparison could not be subjective.

Once information regarding all of the cities was found, questions were written about the cities. The regions were compared to similar regions in the other interface formats. The reason for this was to find cities which were not along the edge or a corner of a region in any of the interfaces. If a city was not along the edge in two interfaces, but was along the edge in the third, it was discarded. While the cities located upon an edge or corner of a region were discarded for questions, when the interface was built information on these cities was still entered so the user could still click on the city.

The other aspect of selecting the questions was the path which each interface would take to get to the city. If a city was chosen from a particular region, all other cities in the same region were discarded for future questions. Since five of the questions would be compound and subjects would have to navigate from one region to another, questions were written so subjects would search for cities not in adjacent regions.

2.4 Using Web-Based Image Maps

Since the interface was designed to be used with a browser, HTML code was created for interaction with the image maps. Each of the maps had a Web-based image map built for it so when subjects selected a certain region the page sent the user to the appropriate page containing the selected region at a higher zoom level. Regions within the image map were created using the rectangular area shape anchor within HTML. Each of the upper level image maps were created using a different formula due to the differing number of regions needing to be created. For example, the 16x4 interface had 16 area shapes at depth equal to one, and four at depth equal to two, while the 8x8 interface had 8 area shapes at all depth levels. Calculations for the rectangular area shapes were based on the methods used in the menu design stage for each of the varying depths.

At the lowest level of depth where cities could be selected, image maps were again used to redirect subjects to the information regarding the selected city. However, unlike the main menus where rectangles were used, 30 pixel circular area shapes were used. This allowed subjects a wider area for selection of the cities. To find the pixel location of the cities on the lowest depth level, Adobe Photoshop [AdobePhotoshop05] was once again used. Each of the images was opened with the rulers turned on and set to display size in pixels. Once the images were opened, a zoom percentage of 500% was used to find the most accurate numbers for the location of the city. The height and depth numbers obtained in Adobe Photoshop [AdobePhotoshop05] were then written into the HTML code for the region.

2.5 Database Design

After the HTML code had been created using the integrated desktop environment (Macromedia Dreamweaver 8 [MacromediaDreamweaver05]), a database was created to store information regarding the interaction of the subjects with the interface. The database was used to store information regarding which region had been selected as well as the timestamps for the transaction. Along with the information about navigation, the database also stored information about whether or not the subject answered the question correctly, the demographic data, and the subject's preferences. Besides subject data, the number of attempts on a particular interface and the number of times the interface was completed was also stored. This would be used in determining which interface a subject used.

Subjects were given an incrementally assigned subject identification number (Subject ID) at the beginning of the first question. Each of the interfaces was given a number, one through six, which was stored with the subject ID in the database. The subject ID was then used in all of the other database tables to tie a subject and interface together to their interaction with the interface.

A table was created for each of the questions and contained the selection the subject made, their subject ID, and the timestamp at which it was made. The naming scheme from the menu design stage (R1, R1_R2, etc.) was used to store the selection in the database. Another table was created to store the subject ID, demographic, and preference information. The final table created stored the subject ID, the question number, whether or not the subject answered the question correctly, and the timestamp when the question was answered. This timestamp was needed to determine the overall question time.

2.6 Dynamic Screen Generation with HTML and Web-Embedded Script Languages

Since the resolution of the monitors was known, the interface was designed in such a way as to prevent horizontal and vertical scrolling in the browser while still displaying the map interface, question, and possible answers. This was done using HTML frames. Since the maximum height of the map had been determined to be 520 pixels and the monitor's resolution was capable of 768 pixels, the other space was used to display a frame with the question and possible answers as well as a zoom out link which sent the subject up one depth level. A vertical frame was constructed with two windows. The top frame contained 240 pixels and was used to display the question and possible answers. The bottom frame used the remaining pixels and contained the map interface and zoom out link.

Part of the experiment was to determine the use of both map and menu labels and their effect on response time and accuracy. For the non-label based interface the link to zoom out needed to be generic while the link for the label based interface needed to contain the name of the continents displayed on the previous depth level. This was accomplished by creating six sets of HTML pages – three for the label based interfaces and three for the non-label based interfaces for each of the varying depths and breadths. For the non-label based interfaces, the link was displayed as: 'Zoom Out' and redirected the subject to one depth level upward. To determine the link text for the label based interface the continents of each of the regions had to be determined. This was done by examining each of the regions and listing each of the continents shown in the image. Once accomplished, the link was displayed to include the continents shown on the interface one depth level upward. As an example, if the previous region showed North America and South America, the link would read: 'Zoom Out to: North America, South America'. For the main menu, or map of the world, the link read: 'Zoom Out to: World'.

PHP was the Web-based script language used to connect to the database because of its ability to interact with MySQL and perform server side functions. The first part of the interface created was the query which identified the interface used by the subject. Subjects only interacted with one interface but enough data needed to be collected for each interface for analysis so the attempts had to be evenly distributed. A database table was created to keep track of subjects who selected the 'Begin' button on the main page or those who started using the interface. The count of each interface was incremented when a subject selected the 'Begin' button. If the subject did not complete the attempt, it was counted as an incomplete attempt and not used in data analysis. At the completion of the questionnaire, the database table was updated to show the number of times the interface had been completed. Completion meant the subject had answered each of the 10 questions as well as the demographic and preference questions. The introduction page to the study contained PHP code which queried the database and retrieved the interface which had been used the least, where least was defined as incomplete attempts plus completed attempts. This method ensured each of the attempts on the interfaces were evenly distributed.

Once a subject began the assignment, their incrementally assigned subject ID was stored in a PHP session variable. Because the test environment was multi-user, storing the subject ID was important so user interaction with the interface could be tracked and assigned to a specific subject. This meant every selection on the map and every question answered needed an association with a subject. Since each of the regions and cities could be selected while a subject interacted with the interface, a way was devised to determine what question the subject was answering for each of the pages. This was done by again taking advantage of a PHP session variable and storing the number of the question the subject was answering.

Hidden form input values and JavaScript were used to redirect subjects when a selection was made using the interface. For example, if a subject selected R2 from the main menu, a form was submitted which redirected the subject to the appropriate page and also posted a variable to the new page indicating the prior page for the subject. When the second page received the posted variable, a PHP script was called to insert a row into the database for the specific selection, user, and question. This information was later used during the data analysis phase. Hidden forms were submitted via the OnClick JavaScript command. Each of the regions of the map contained an anchor reference to a null anchor ('#') as well as an OnClick method for the region selected. For example, if R2_R1 was selected from R2, the JavaScript command would be: OnClick = 'document.r2.submit();' where R2 was a form written in the page containing a hidden text input type with a value of 'R2'. This form redirected the subject to the corresponding region they selected and the PHP script would insert the 'R2' selection into the database with a corresponding timestamp.

When a subject answered a question by selecting the appropriate answer from the radio list and selecting the 'Continue' button, the answer selected was posted to a page containing the next question and a 'Begin' button on an otherwise blank screen. Since each of the subjects answered the same 10 questions in the same order, the answers to the questions were hard coded inside PHP code. When the answer was posted to the page, it was determined whether the answer was correct or incorrect. A value of 'Y' or 'N' was then stored in the database based on the correctness of the answer selected. The value was stored with its corresponding timestamp which was used to determine the overall time spent on the question.

Individual pages for the cities were then designed. Since the cities were able to be selected from multiple regions based on which interface was being used, multiple versions of the PHP pages containing city information had to be built. Within each of these pages were two facts about the city: the population and a random fact. Each city page also contained a link to go back to the map interface. This was needed for subjects who chose the wrong city to zoom back out as well as when answering compound questions which required finding information about multiple cities. Since each of the interfaces had a different region for the city, each version of the city pages contained a link to the appropriate region for each interface.

Once all of the interfaces were built and the PHP scripts were written, testing was completed to ensure they worked properly. The first test activity was to find every city for which information was gathered, this included cities that were part of the questions to be answered as well as those cities that were not part of any question, for each of the interfaces. During the process of finding the cities, interaction with the interface was tested to ensure when a subject chose the correct region they were redirected to the correct page. Once these two tests had been completed, regions containing no city information were also tested. These areas included Antarctica as well as sections of the Atlantic and Pacific Oceans.

As testing was taking place on the interface, the database was also tested. A series of designated test runs was designed to verify the database was storing the correct regions selected, answers, as well as the demographic and preference questions. The tests were run on a predetermined path for each of the interfaces. The database was then queried to verify each of the test runs. At the completion of the database verification, the system was made available for subjects to use. During subject use, the database was routinely backed up and saved to another computer to ensure little or no data would be lost in the case of a catastrophic server failure.

To prevent subjects from being able to continue to the next question without answering the current question, code was created that forced subjects to answer each of the questions before moving onto the next question. If a subject selected the 'Continue' button for a question without first selecting a radio button from the list, an error message was displayed to notify the suhject. This was also done for the demographic and preference questions to ensure complete data was collected for each of the subjects. If a subject attempted to move to the next question without answering the current question, it was not counted as an error for the question.

2.7 Web Server Information

An HTTP server was used on a host machine running SunOS 5.10. Apache was chosen as the HTTP Web server and was built from source code to ensure the libraries would not be shared with other Apache instances running on the same server. PHP was built from source code for use with Apache with the same constraints. A user was created on the server and given limited access to server resources. Apache and PHP were downloaded to the server where PHP was installed as an embedded static module for Apache by following the tutorial found on the PHP website. Extra configuration directives were needed to install all appropriate files in the user's home directory. PHP also had separate configuration directives to ensure compatibility with MySQL. After completion of the installation, Apache was configured to receive requests on port 8003.

2.8 Subjects

Subjects in the study were junior college students enrolled in algebra-based mathematics courses. Information was gathered about students in the study; however subjects remained anonymous. Information gathered included: gender, eyesight, educational level, computer experience, age, and handedness.

2.9 Completion of the Experiment

In order to prevent skewing of results based on machine processing speed, Internet connection, or monitor resolution each of the experiments were done on the same set of machines at the same location. Each of the machines had the same hardware specifications, monitor resolution, and Internet connection. Subjects were asked to use a randomly assigned interface to seek information on a city, or a comparison of two cities.

2.10 Data Gathering and Compilation

Subjects were given a two week period to voluntarily participate in the experiment. Subjects used computers set up in the same physical location with the same Internet connection speed. Subjects had the ability to adjust seat and LCD monitor height for individual comfort as well as being able to adjust the monitor's brightness and contrast if desired. During data gathering the database and server were frequently monitored for system errors and nightly database backups were stored off-site. The database was queried to determine how many subjects had started and completed the experiment for each of the interfaces. At the end of the two week period, 89 subjects had completed the assessments and each interface had at least 13 valid attempts.

2.10.1 Data Filtering

Validity of the attempts was determined by exporting the data from the database where data filtering could take place. A query was used to recover all of the data for subjects who completed the experiment by checking a flag stored in each of their records which was modified when the subject completed the demographic and preference questions.

Once the subject ID's were found, each of the individual question tables were queried for the completed subjects and compiled into a master data set where filters were used to look for transactions made for an individual subject ID for all 10 questions. Since the site was available via the World Wide Web, people outside of the subject group were able to access the site. Internet Protocol (IP) addresses for each of the attempts were stored in the database when demographic data was gathered. Data from IP addresses outside the range of the computer lab was discarded. Data from any subject who did not complete the experiment was discarded. Subjects who did not reach the correct region before answering the question were also discarded as invalid attempts. The remaining data was used to determine response times, accuracy, and preferences.

2.10.2 Computing of Metrics for Factorial Analysis

Figure 9 graphically shows the factorial research framework for the study. Three independent variables were identified: interface format, use of labels, and question complexity. Between-subjects statistical analysis was completed for interface format and use of labels and within subjects analysis was completed for question complexity. Analysis of variance (ANOVA) between groups was performed on data that was within three standard deviations from the mean. Seven data records were identified as outliers and removed from the data set before ANOVA calculations were completed. Response time and accuracy metrics were then computed. Overall question time was determined by running a database query to obtain the start time for each of the questions for each of the valid subject ID's. Another query was then run to obtain the time the question was answered. Overall question time was defined as the number of seconds between the two times. After the overall question time was computed and stored, the initial response time was computed using the same start time. A query was run to obtain the first transaction in the database after the subject began the question. Initial response time was defined as the number of seconds between these two times and its value was also stored.

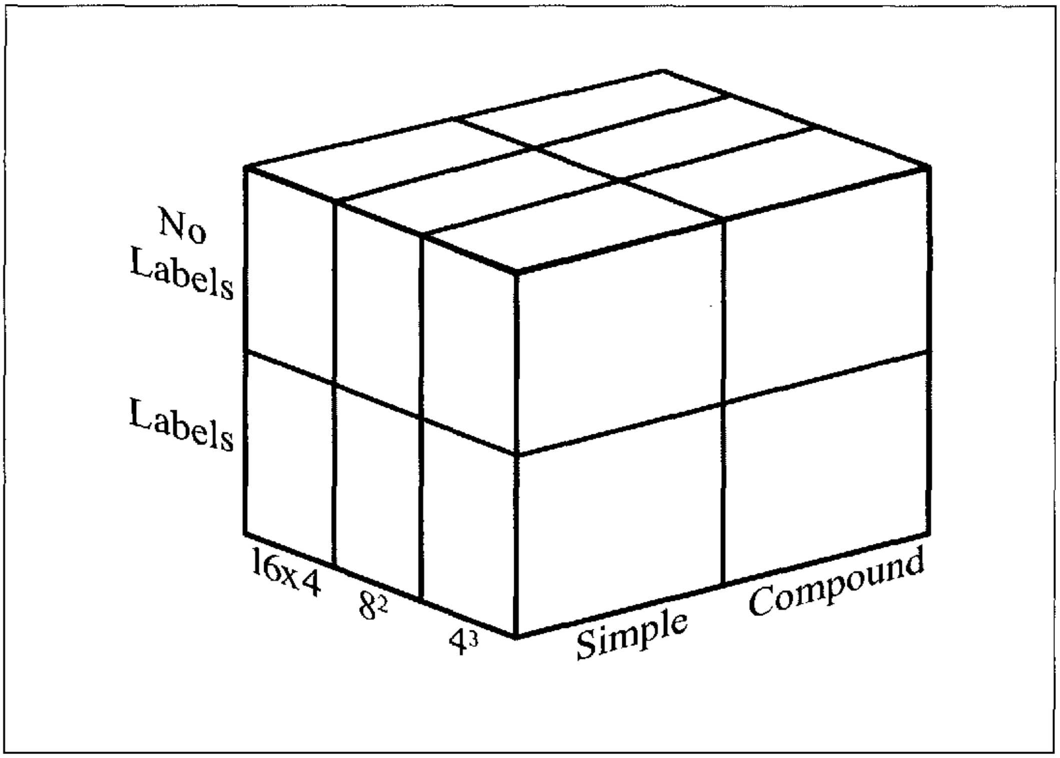

Figure 9: Factorial Research Framework

After the response time metrics were calculated, the accuracy metrics were determined. The answer table was queried with the valid subject ID's to retrieve the subject's answers to the 10 questions. To determine the number of errors for each subject, each question was examined individually. For the simple questions, the subject's interaction was compared to the ideal path the subject could have taken to retrieve the information. Using the definition of an error as a selection of the wrong region, the number of errors for each simple question was determined one subject at a time. Finding the number of errors for the compound questions followed a similar methodology. The compound questions differed because there were two separate ideal paths the subject could have taken depending on which city was chosen first to find. The subject's path was compared to these paths to determine the number of errors for the compound questions. The number of errors for simple and compound questions were then added to the data set. The total number of selections the user made with the interface was also stored.

The final information integrated into the research data set was the preference and demographic data. The table was queried and the values were stored for each of the valid subject ID's. Once the data was collected and stored in a consistent manner, statistical analysis was done.

Chapter 3

RESULTS

Multiple types of statistics were calculated from the data set including: frequencies, crosstabs, correlations, t-tests, and analysis of variations. Summary and descriptive statistics were generated to verify the validity of the data collected. During this step an obvious data anomaly was found. One data point for the initial response time for the second simple question was found to be negative. That single value was marked as missing and the remaining statistical analyses was performed.

3.1 Subjects

Frequencies and crosstabs were calculated to determine the general characteristics of the subject group. The gender distribution of the 89 subjects who completed the study was 58 female and 31 male. The majority of subjects were right handed (81) and had 20/20 vision either naturally or corrected (58). Most of the students considered their level of computer experience as intermediate (55) or advanced (26). For a full list of subject frequency results see Appendix D.

Frequencies for interface type were also calculated. Table 1 shows the frequencies for each of the interface formats and Table 2 shows the frequencies for the label type. These statistics were complimented by calculating crosstabs for the subject group and the interface types. Crosstabs were calculated for handedness, eyesight, and computer experience. Results showed more females (37) completed the study with label-based menus than without (21). Conversely, fewer males completed the label-based study (8) than without (23). These differences were not statistically significant.

Format		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4x4x4	30	33.7	33.7	33.7
	8x8	30	33.7	33.7	67.4
	16x4	29	32.6	32.6	100.0
	Total	89	100.0	100.0	

Table 1: Interface Format Frequencies

Label		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	With	45	50.6	50.6	50.6
	Without	44	49.4	49.4	100.0
	Total	89	100.0	100.0	

Table 2: Label and Non-label Frequencies

For the 37 females who completed the label-based interface, 13 used the 4^3 interface, 13 used the 8^2 interface, and 11 used the 16x4 interface. The results for females who used the non-label based interfaces were similarly spread out where five, nine, and seven completed the 4^3, 8^2, and 16x4 interfaces respectively. For the eight males who completed the label based study two used the 4^2 interface, four used the 8^2 interface, and two used the 16x4 interface. For the 23 males who completed the non-label based study ten, four, and nine completed the 4^3, 8^2, and 16x4 interfaces respectively. Crosstabs for handedness, eyesight, and computer experience showed similar balanced results. For a full list of subject group crosstabs see Appendix E. The differences between the handedness, eyesight, and computer experience of the subject groups were not statistically significant.

3.2 Correlations

After the completion of the summary and descriptive statistics, correlations were calculated. Some of the significant correlations were obvious, such as a relationship among user activity, or clicks, and errors on the same question. These correlations were then examined and logical groupings were made. Other correlations were identified as significant when the p value was less than 0.05 and the absolute value of r was greater than 0.30. For a full list of significant correlations see Appendix F.

3.2.1 Accuracy

One of the logical groups found in the data was the accuracy – based on user activity – within the interface. There were several significant correlations for user activity among question types. For example, user activity on the first simple question had a relationship with the amount of activity on the second simple question ($p < 0.01$, $r = 0.635$), the third simple question ($p < 0.01$, $r = 0.327$), and the fifth simple question ($p < 0.01$, $r = .347$). Similarly the second simple question had a relationship with the third ($p < 0.01$, $r = 0.476$) and fifth ($p < 0.01$, $r = 0.569$) simple question. There were also relationships among the fourth and fifth ($p < 0.01$, $r = 0.435$) simple questions. Similarly, there was a relationship among compound questions where the amount of user activity on the first complex question had a relationship with the second ($p < 0.01$, $r = 0.487$), third ($p < 0.01$, $r = 0.663$), and fourth ($p < 0.01$, $r = 0.335$) compound questions. There were significant correlations for the second compound question with the third, fourth and fifth compound questions, the third with the fourth compound question, and the fourth with the fifth compound question. Table 3 shows a full listing of significant correlation values for user activity among question types.

Correlation Pair		r	p
q1 clicks simple	q2 clicks simple	0.635	0.000
q1 clicks simple	q3 clicks simple	0.327	0.002
q1 clicks simple	q5 clicks simple	0.347	0.001
q2 clicks simple	q3 clicks simple	0.476	0.000
q2 clicks simple	q5 clicks simple	0.569	0.000
q4 clicks simple	q5 clicks simple	0.435	0.000
q1 clicks compound	q2 clicks compound	0.487	0.000
q1 clicks compound	q3 clicks compound	0.663	0.000
q1 clicks compound	q4 clicks compound	0.335	0.001
q2 clicks compound	q3 clicks compound	0.337	0.001
q2 clicks compound	q4 clicks compound	0.514	0.000
q2 clicks compound	q5 clicks compound	0.537	0.000
q3 clicks compound	q4 clicks compound	0.412	0.000
q4 clicks compound	q5 clicks compound	0.677	0.000

Table 3: Correlation Values for User Activity among Question Types

There was a similar significant correlation between user activity and errors among question types. The amount of user activity on the first simple question had a relationship with the number of errors on the second ($p < 0.01$, $r = 0.619$), third ($p < 0.01$, $r = 0.324$), and fifth ($p < 0.01$, $r = 0.341$) simple questions. Likewise, there was a relationship among the number of errors committed on the first simple question and the amount of user activity on the second ($p < 0.01$, $r = 0.652$), third ($p < 0.01$, $r = 0.320$), and fifth ($p < 0.01$, $r = 0.365$) simple questions. There was a similar relationship among the compound questions where the amount of user activity on the first compound question had a relationship with the number of errors committed on the second ($p < 0.01$, $r = 0.509$), third ($p < 0.01$, $r = 0.611$), and fourth ($p < 0.01$, $r = 0.338$) compound questions. Table 4 contains a full list of significant correlation values for user activity and errors for simple questions and Table 5 contains the full list for compound questions.

Correlation Pair		r	p
q1 clicks simple	q2 errors simple	0.619	0.000
q1 clicks simple	q3 errors simple	0.324	0.002
q1 clicks simple	q5 errors simple	0.341	0.001
q1 errors simple	q2 clicks simple	0.652	0.000
q1 errors simple	q3 clicks simple	0.320	0.002
q1 errors simple	q5 clicks simple	0.365	0.000
q2 clicks simple	q3 errors simple	0.485	0.000
q2 clicks simple	q5 errors simple	0.567	0.000
q2 errors simple	q3 clicks simple	0.456	0.000
q2 errors simple	q5 clicks simple	0.568	0.000
q4 clicks simple	q5 errors simple	0.439	0.000
q4 errors simple	q5 clicks simple	0.378	0.000

Table 4: Correlation Values for User Activity
and Errors among Simple Questions

The fifth compound question had the most relationships with other question types. User activity for the first ($p < 0.01$, $r = -0.561$), second ($p < 0.01$, $r = -0.740$), third ($p < 0.01$, $r = -0.443$), and fifth ($p < 0.01$, $r = -0.477$) simple questions were all related to the fifth compound question. Since user activity and number of errors were closely tied, it was logical that the incorrect answers to the fifth compound question had the most relationships with the number of errors on simple questions. Table 6 contains a full list of significant relationships among user activity and compound questions.

Correlation Pair		r	p
q1 clicks compound	q2 errors compound	0.509	0.000
q1 clicks compound	q3 errors compound	0.611	0.000
q1 clicks compound	q4 errors compound	0.338	0.001
q1 errors compound	q2 clicks compound	0.468	0.000
q1 errors compound	q3 clicks compound	0.666	0.000
q1 errors compound	q4 clicks compound	0.319	0.002
q2 clicks compound	q3 errors compound	0.325	0.002
q2 clicks compound	q4 errors compound	0.523	0.000
q2 clicks compound	q5 errors compound	0.570	0.000
q2 errors compound	q3 clicks compound	0.362	0.000
q2 errors compound	q4 clicks compound	0.508	0.000
q2 errors compound	q5 clicks compound	0.493	0.000
q3 clicks compound	q4 errors compound	0.412	0.000
q3 errors compound	q4 clicks compound	0.392	0.000
q4 clicks compound	q5 errors compound	0.677	0.000
q4 errors compound	q5 clicks compound	0.650	0.000

Table 5: Correlation Values for User Activity
and Errors among Compound Questions

Another relationship found within the correlation results for accuracy was for overall response time to incorrect answers. Again, the fifth compound question was related to simple question types. In this instance, there was a statistically significant relationship with an incorrect answer on the fifth question to the overall response time for all of the simple questions. Table 7 shows the significant correlation values for overall response time to incorrect answers.

Correlation Pair		r	p
q2 clicks simple	q1 answer correctness compound	-0.302	0.004
q3 clicks simple	q1 answer correctness compound	-0.376	0.000
q5 clicks simple	q3 answer correctness compound	-0.302	0.004
q1 clicks simple	q4 answer correctness compound	-0.315	0.003
q2 clicks simple	q4 answer correctness compound	-0.317	0.002
q3 clicks simple	q4 answer correctness compound	-0.366	0.000
q1 clicks simple	q5 answer correctness compound	-0.561	0.000
q2 clicks simple	q5 answer correctness compound	-0.740	0.000
q3 clicks simple	q5 answer correctness compound	-0.443	0.000
q5 clicks simple	q5 answer correctness compound	-0.477	0.000
q3 errors simple	q1 answer correctness compound	-0.372	0.000
q1 errors simple	q4 answer correctness compound	-0.309	0.003
q3 errors simple	q4 answer correctness compound	-0.348	0.001
q1 errors simple	q5 answer correctness compound	-0.562	0.000
q2 errors simple	q5 answer correctness compound	-0.718	0.000
q3 errors simple	q5 answer correctness compound	-0.442	0.000

Table 6: Correlation Values for User Activity
and Incorrect Compound Answers

Correlation Pair		r	p
q2 ov resp time simple	q1 answer correctness compound	-0.325	0.002
q3 ov resp time simple	q1 answer correctness compound	-0.318	0.002
q2 ov resp time compound	q1 answer correctness compound	-0.317	0.002
q2 ov resp time simple	q4 answer correctness compound	-0.328	0.002
q1 ov resp time simple	q5 answer correctness compound	-0.450	0.000
q2 ov resp time simple	q5 answer correctness compound	-0.753	0.000
q3 ov resp time simple	q5 answer correctness compound	-0.380	0.000
q4 ov resp time simple	q5 answer correctness compound	-0.484	0.000
q5 ov resp time simple	q5 answer correctness compound	-0.476	0.000

Table 7: Correlation Values for Overall Response Time
and Incorrect Compound Answers

3.2.2 Response Time

Another logical group found in the correlation statistics dealt with overall response time among question types. One obvious significant correlation was the amount of user activity versus the overall response time for a question. However, relationships found among user activity and overall response time were for questions of the same type. For example, the amount of user activity on the first simple question had a relationship with the overall response time for the second ($p < 0.01$, $r = 0.641$), fourth ($p < 0.01$, $r = 0.318$), and fifth ($p < 0.01$, $r = 0.323$) simple questions. The overall response time for the first simple question had a relationship with the amount of user activity on the second ($p < 0.01$, $r = 0.518$) and fifth ($p < 0.01$, $r = 0.346$) simple questions. This type of relationship was also found within the compound question types where the first compound question had a relationship with the overall response time for the third ($p < 0.01$, $r = 0.474$) compound question. The amount of user activity on the second compound question had relationships with the overall response time on the third ($p < 0.01$, $r = 0.357$), fourth ($p < 0.01$, $r = 0.432$), and fifth ($p < 0.01$, $r = 0.323$) compound questions. Table 8 contains a full list of significant correlation values for this group.

Overall response time had a relationship with the number of navigation errors committed on similar question types. For example, the overall response time for the first simple question was related to the number of errors committed on the second ($p < 0.01$, $r = 0.499$) and fifth ($p < 0.01$, $r = 0.337$) simple questions. The number of errors committed on the first simple question was likewise related to the overall response time for the second ($p < 0.01$, $r = 0.662$), fourth ($p < 0.01$, $r = 0.345$), and fifth ($p < 0.01$, $r = 0.343$)

simple questions. This relationship held true for compound questions where the overall response time for the third compound question was related to the number of errors committed on the first ($p < 0.01$, $r = 0.471$), second ($p < 0.01$, $r = 0.399$), and fourth ($p < 0.01$, $r = 0.347$) compound questions. A full list of significant correlations values for this group is shown in Table 9.

Correlation Pair		r	p
q1 clicks simple	q2 ov resp time simple	0.641	0.000
q1 clicks simple	q4 ov resp time simple	0.318	0.002
q1 clicks simple	q5 ov resp time simple	0.323	0.002
q1 ov resp time simple	q2 clicks simple	0.518	0.000
q1 ov resp time simple	q5 clicks simple	0.346	0.000
q2 clicks simple	q3 ov resp time simple	0.405	0.000
q2 clicks simple	q4 ov resp time simple	0.505	0.000
q2 clicks simple	q5 ov resp time simple	0.541	0.000
q2 ov resp time simple	q3 clicks simple	0.492	0.000
q2 ov resp time simple	q4 clicks simple	0.337	0.001
q2 ov resp time simple	q5 clicks simple	0.557	0.000
q4 clicks simple	q5 ov resp time simple	0.431	0.000
q4 ov resp time simple	q5 clicks simple	0.646	0.000
q1 clicks compound	q3 ov resp time compound	0.474	0.000
q2 clicks compound	q3 ov resp time compound	0.357	0.010
q2 clicks compound	q4 ov resp time compound	0.432	0.000
q2 clicks compound	q5 ov resp time compound	0.323	0.002
q2 ov resp time compound	q4 clicks compound	0.390	0.000
q3 clicks compound	q4 ov resp time compound	0.376	0.000
q3 ov resp time compound	q4 clicks compound	0.401	0.000
q4 clicks compound	q5 ov resp time compound	0.469	0.000
q4 ov resp time compound	q5 clicks compound	0.398	0.000

Table 8: Correlation Values for Overall Response Time and User Activity

Correlation Pair		r	p
q1 errors simple	q1 ov resp time simple	0.578	0.000
q1 errors simple	q2 ov resp time simple	0.662	0.000
q1 errors simple	q4 ov resp time simple	0.345	0.001
q1 errors simple	q5 ov resp time simple	0.343	0.001
q1 ov resp time simple	q2 errors simple	0.499	0.000
q1 ov resp time simple	q5 errors simple	0.337	0.001
q2 errors simple	q3 ov resp time simple	0.389	0.000
q2 errors simple	q5 ov resp time simple	0.541	0.000
q2 ov resp time simple	q3 errors simple	0.505	0.000
q2 ov resp time simple	q4 errors simple	0.322	0.002
q2 ov resp time simple	q5 errors simple	0.556	0.000
q4 errors simple	q5 ov resp time simple	0.386	0.000
q4 ov resp time simple	q5 errors simple	0.653	0.000
q1 errors compound	q3 ov resp time compound	0.471	0.000
q2 errors compound	q3 ov resp time compound	0.399	0.000
q2 errors compound	q4 ov resp time compound	0.461	0.000
q2 errors compound	q5 ov resp time compound	0.331	0.002
q2 ov resp time compound	q4 errors compound	0.411	0.000
q2 ov resp time compound	q5 errors compound	0.322	0.002
q3 errors compound	q4 ov resp time compound	0.347	0.001
q3 ov resp time compound	q4 errors compound	0.414	0.000
q4 errors compound	q5 ov resp time compound	0.456	0.000
q4 ov resp time compound	q5 errors compound	0.407	0.000

Table 9: Correlation Values for Overall
Response Time and Number of Errors

3.2.3 Preferences

The preference questions were included in the correlation statistics and were grouped into two categories based on user activity and number of errors committed for each question type. For the simple questions, the user activity on the third question had the most relationships with preference questions. A full list of preference questions is available in Appendix C. The user activity for the third simple question had a relationship with the first ($p < 0.01$, $r = -0.399$), third ($p < 0.01$, $r = -0.345$), and sixth ($p < 0.01$, $r = -0.315$) preference questions. The user activity for the third compound question and the total amount of user activity for compound questions had the most relationships with preference questions. The user activity for the third compound question had a relationship with the first ($p < 0.01$, $r = -0.446$), third ($p < 0.01$, $r = -0.459$), and fifth ($p < 0.01$, $r = 0.352$) preference questions. The total amount of user activity for compound questions was related to the first ($p < 0.01$, $r = -0.391$), second ($p < 0.01$, $r = -0.331$), and third ($p < 0.01$, $r = -0.384$) preference questions. Table 10 contains the full list of significant correlation values for preferences and user activity.

Correlation Pair		r	p
q1 clicks simple	second preference question	-0.359	0.001
q2 clicks simple	second preference question	-0.471	0.000
q3 clicks simple	first preference question	-0.399	0.000
q3 clicks simple	third preference question	-0.345	0.001
q3 clicks simple	sixth preference question	-0.315	0.003
q4 clicks simple	fourth preference question	-0.332	0.001
q5 clicks simple	second preference question	-0.360	0.001
first preference question	total clicks simple	-0.309	0.000
second preference question	total clicks simple	-0.470	0.000
q1 clicks compound	first preference question	-0.407	0.000
q3 clicks compound	first preference question	-0.446	0.000
q3 clicks compound	third preference question	-0.459	0.000
q3 clicks compound	fifth preference question	-0.352	0.001
q4 clicks compound	second preference question	-0.368	0.000
first preference question	total clicks compound	-0.391	0.000
second preference question	total clicks compound	-0.331	0.002
third preference question	total clicks compound	-0.384	0.000

Table 10: Correlation Values for Preferences and User Activity

Like response time, preferences also had relationships with the number of navigation errors committed for the two types of questions. The second preference statement – "It was easy to find the information I was looking for" – was related to three of the simple questions: the first ($p < 0.01$, $r = -0.358$), second ($p < 0.01$, $r = -0.458$), and fifth ($p < 0.01$, $r = -0.352$). The second preference question was also related to the total number of errors committed on simple questions ($p < 0.01$, $r = -0.462$). The number of errors committed on the third compound question had relationships with three of the preference questions: the first ($p < 0.01$, $r = -0.474$), third ($p < 0.01$, $r = -0.475$), and fifth ($p < 0.01$, $r = -0.386$). Table 11 shows the significant correlation values for this group.

Correlation Pair		r	p
q1 errors simple	second preference question	-0.358	0.001
q2 errors simple	second preference question	-0.458	0.000
q3 errors simple	third preference question	-0.324	0.002
q4 errors simple	fourth preference question	-0.359	0.001
q4 errors simple	fifth preference question	-0.330	0.002
q5 errors simple	second preference question	-0.352	0.001
second preference question	total errors simple	-0.462	0.000
q1 errors compound	first preference question	-0.380	0.000
q3 errors compound	first preference question	-0.474	0.000
q3 errors compound	third preference question	-0.475	0.000
q3 errors compound	fifth preference question	-0.386	0.000
q4 errors compound	second preference question	-0.364	0.000
first preference question	total errors compound	-0.392	0.000
second preference question	total errors compound	-0.339	0.001
third preference question	total errors compound	-0.396	0.000

Table 11: Correlation Values for Preferences and Errors Committed

3.2.4 Other Correlations

There were other significant correlations which consisted of smaller logical groups and results where no grouping could occur. One such result was the number of errors committed and its relationship with errors committed for other questions of the same type. The number of errors for the first simple question showed a relationship with the number of errors for the second ($p < 0.01$, $r = 0.641$), third ($p < 0.01$, $r = 0.325$), and fifth ($p < 0.01$, $r = 0.365$) simple questions. For compound questions, the number of errors on the second question was related to the number of errors for the third ($p < 0.01$, $r = 0.350$), fourth ($p < 0.01$, $r = 0.527$), and fifth ($p < 0.01$, $r = 0.540$) compound questions.

Results for overall response time also showed relationships to overall response time in other questions of the same type. For example, the overall response time for the first simple question was related to the overall response time for the second ($p < 0.01$, $r = 0.530$) and fifth ($p < 0.01$, $r = 0.312$) simple questions. The second compound question's overall response time was related to the overall response time of the third ($p < 0.01$, $r = 0.425$), fourth ($p < 0.01$, $r = 0.511$), and fifth ($p < 0.01$, $r = 0.370$) compound questions. The full list of significant correlations is available in Appendix F.

3.3 Analysis of Variance

Several analysis of variance (ANOVA) calculations were performed on total clicks, navigational errors, incorrect answers, and overall response time for simple and compound questions, as well as on the grand totals for each of the same categories over three main effects: presence of labels, interface format, and type of question. Post hoc testing was performed using the Least Significant Difference (LSD) test on any statistically significant ANOVA result.

The interaction of question type and presence or absence of labels had an effect on the total number of errors committed for simple and compound questions ($p < 0.5$). Table 12 shows the results of this effect. Analyzing the means of the number of errors for simple and compound questions showed the interaction of simple questions with label based interfaces had statistically significantly fewer errors than simple questions with non-label based interfaces as well as compound questions with or without labels. The means calculations also showed significantly fewer errors were committed using the non-label based interface when answering compound questions than when the label based interface was used. Appendix G shows the means calculations for simple and compound questions in label and non-label based interfaces. The presence or absence of labels also had an effect on the total errors committed ($p < 0.01$). However, overall means calculations based on question type showed more errors were committed for compound questions. Results of the means calculations for simple and compound questions are available in Appendix H.

Source	SS	df	MS	F	p
qtype	854.825	1	854.825	11.518	0.001
qtype * label	363.946	1	363.946	4.904	0.030
Error	5640.350	76	74.215		

Table 12: ANOVA Results for Total Errors Committed
on Simple and Compound Questions

The same combination of question type and presence of labels had an effect on the overall response time for simple and compound questions (p < 0.01). Results of the calculations for overall response time can be seen in Table 13. After the means calculations for overall response time were completed, the results show response time was significantly less when responding to simple questions in label based interfaces. Like the results for errors, means calculations showed overall response time for compound questions was less when using non-label based interfaces than in label based interfaces. Results of the means calculations can be found in Appendix G. The question type alone had an effect on the overall response time taken for simple and compound questions (p < 0.01). These results are shown in Table 13. Means calculations for this result showed the expected result of significantly less time on simple questions. Appendix H contains the results of these calculations. Other ANOVA calculations on overall response time showed that while question type alone, or in conjunction with labels, had significant results, the presence or absence of labels alone had an effect (p < 0.05). Table 14 contains these results. Means calculations for this result showed there was significantly less overall response time for subjects who used the non-label based interface. The results of these calculations can be found in Appendix I.

Source	SS	df	MS	F	p
qtype	2909177.488	1	2909177.488	57.510	0.000
qtype * label	400461.521	1	400461.521	7.917	0.006
Error	3844480.682	76	50585.272		

Table 13: ANOVA Results for Total Time on Simple and Compound Questions

Source	SS	df	MS	F	p
label	277218.156	1	277218.156	4.174	0.045
Error	5047847.441	76	66419.045		

Table 14: ANOVA Results for Labels on Total Time
for Simple and Compound Questions

The next ANOVA calculation was done on the effect of question type in label and non-label interfaces on the number of clicks. Two significant results were found. The first result, as seen in Table 15, showed the type of question had an effect on the total number of clicks. Means calculations showed simple questions had significantly less clicks than compound questions. Means calculation results can be found in Appendix H. The second result, also shown in Table 15, showed the interaction of question type and the presence or absence of labels had an effect on the number of clicks. Means calculations, as seen in Appendix G, showed the least number of clicks was made while subjects were answering simple questions in a label based interface. Results showed there were significantly fewer clicks when subjects were using the non-label based interface answering compound questions than when subjects used the label based interface.

Source	SS	df	MS	F	p
qtype	69378.323	1	69378.323	207.182	0.000
qtype * label	1477.415	1	1477.415	4.412	0.039
Error	25449.840	76	334.866		

Table 15: ANOVA Results for Total Clicks
on Simple and Compound Questions

The final result from the ANOVA calculations was found with respect to the number of incorrect answers. Question type alone was found to have an effect on the total incorrect answers for simple and compound questions ($p < 0.01$). Results of the ANOVA calculation are shown in Table 16. Means calculations, found in Appendix H, showed there were fewer correct answers made on compound questions over simple questions. More importantly, the three factors of question type, presence or absence of labels, and interface format all interacted with the total incorrect answers for simple and compound questions ($p < 0.05$). Results from the calculation of the ANOVA are found in Table 16. Means calculations done on this result showed the number of correct answers was significantly fewer in the 4^3, non-label based interface when answering compound questions. Results of this calculation can be found in Appendix J.

Source	SS	df	MS	F	p
qtype	0.743	1	0.743	10.972	0.001
qtype * label * iform	0.454	2	0.227	3.351	0.040
Error	5.148	76	0.068		

Table 16: ANOVA Results for Task Accuracy
on Simple and Compound Questions

Chapter 4

DISCUSSION

Empirical data was gathered from systems implemented based on prior research and results for the issues of scent, depth versus breadth, and task complexity were examined. Results suggested subjects used different strategies when faced with tasks of varying complexity. In a result counter to previous research completed, the interface format was found not to have a statistically significant impact on the subject's results. Subjects also performed more efficiently overall when using the non-label based interface for compound questions; however when using the 4^3, non-label based interface, subjects were less accurate when answering compound questions. A detailed discussion of the results follows.

When the correlation data was examined, user activity was related to the accuracy of the question answered. There were some relationships among user activity and response time to task accuracy for simple questions. However, based on the frequency of the number of subjects who answered the simple questions incorrectly these relationships do not offer deeper insight into the data. Accuracy of the question answered did cross over from one question type to another when it was related to user activity, errors, and overall response time. This result implied accuracy of the answer may be based on user activity and the overall response time.

Preferences were the only metrics not measured automatically within the interface. However, the significant correlations showed subjects did not respond the same to each of the preference questions. Furthermore, the negative correlation regarding user activity and errors committed showed subjects who performed worse were less likely to agree with the preference questions – which were all stated in a positive manner. For example, the subjects with higher amounts of user activity for the first, second, and fifth simple questions agreed less to the second preference question – "I was never at a loss to what the next step was" – than those who had less user activity. Relationships also existed among user activity in compound questions to preference questions. These results suggested when subjects were confused about the next step to take using the interface they committed more navigation errors.

Analysis of variance calculations showed varying interactions of question type, interface format, and presence or absence of labels on accuracy and response time within the data. Many of the results were to be expected. For example, there were significantly more errors committed on compound questions than simple questions. Overall response time, number of clicks, and number of incorrect answers were also less when responding to simple questions.

4.1 Significant Findings

The correlation statistics showed interesting results regarding accuracy, response time, and preferences. While there was a relationship among simple questions to other simple questions as well as compound to compound regarding user activity and number of errors committed, there was no relationship found for simple to compound questions except for overall response time to incorrect compound answers. Response times for simple questions were related to user activity on other simple questions. The same response time relationship was true for compound to compound questions. One possible explanation of the correlation results was that subjects used different solution strategies to answer simple and compound questions. Many relationships were found among simple to simple, or compound to compound questions, yet only one statistically significant result was found relating simple to compound questions. A possible explanation for the lack of relationships between simple and compound questions was that subjects preferred the close-ended structure of the compound questions to the open-ended structure of the simple questions. The subjects may have been task oriented in nature and were able to use the interface to better follow the compound line of questioning.

The interface format, or its depth and breadth, was found to not have a statistically significant impact on the results. This result was counter to previous research completed in text-based hierarchical menu systems where the results generally favored broader, shallower menu structures. In the implementation of the multidimensional graphical map-based menu system, the varying depths and breadths had no significant impact on accuracy, response time, or preferences. These results suggested design techniques based on text-based hierarchical menu systems may not be generalized to implementation of a multidimensional graphical menu system or the depth and breadth of the interface needed to significantly increase in order to see similar results.

Based on the ANOVA calculations there were significant findings with respect to errors committed. While there were significantly fewer errors committed when responding to simple questions as opposed to compound, subjects made fewer errors when using the non-label based interfaces to respond to compound questions. This result was duplicated in total number of clicks. While there were significantly less clicks made when responding to simple questions, as well as when the label based interface was used, subjects made fewer clicks when using the non-label based interface to answer compound questions. The same result can also be seen with respect to overall response time. Again, the overall response time was less when responding to simple questions, however, subjects answered compound questions in significantly less time when using the non-label based interface. Furthermore, the total mean overall response time was significantly less when non-label based interfaces were used. These results showed the non-label based interface was more efficient than the label based interface for compound questions. One possible explanation for this result was subjects intuitively used the graphical format of the interface. The presence of labels may have been seen more as a distraction than an aid in answering compound questions.

The final significant result found in the ANOVA calculations was the interaction of the three main factors of interface format, question type, and presence or absence of labels on the number of incorrect answers. While the results showed there were significantly fewer incorrect answers when responding to simple rather than compound questions, means calculations showed the interaction of the three main factors was most significant with respect to the 4^3, non-label based interface when responding to compound questions. The results showed this combination provided significantly more errors than any other combination. This result can also be seen in Gray's study of hierarchical menu systems where subjects made more errors at greater depth levels without the use of menu titles [Gray86]. Gray's results suggested as depth levels increase, the use of menu titles and the presence of labels on the multidimensional map interface make the subject treat the interface more like a text-based interface and less like a graphical interface. It is possible subjects relied more on the scan-and-match process that Landauer described [Landauer85] when presented with the 4^3, label based interface; however, subjects became more confused at greater depth levels using the 4^3, non-label based interface.

The results from the ANOVA calculations showed interface format had no statistically significant impact on accuracy, response time, or preferences. Other results showed while subjects were more efficient overall using the non-label based interface for compound questions, they answered significantly more questions incorrect when using the 4^3, non-label based interface. These results suggested the combination of the greater depth level of the 4^3 interface and the absence of labels resulted in more confusion in subject responses. It also suggested at greater depth levels the additional textual information of labels on the map was useful when answering compound questions. Overall, while answering compound questions using the non-label based interface, subjects were more efficient at using the graphical representation of the map and committed fewer errors, made less clicks, and completed the question in less time than subjects answering compound questions using the label based interface. This suggested an interface combining textual and graphical information would be more accurate with deeper interfaces while the graphical information alone would be more accurate with broad interfaces – particularly with complex tasks.

4.2 Future Work

There are many areas of future work which could be performed to extend the results of this study. The results of this study showed no significant difference in measured variables among the interface types, therefore one possible extension of this study would be to create significantly deeper and broader trees for which to traverse. The breadth and depth of the study could easily be increased by the use of Web-based image maps and a smaller zooming factor.

Secondly, the introduction of a three-dimensional graphical representation of the map may be studied. Some Graphical Information Systems use Global Positioning System data to display latitudinal and longitudinal data on topographical maps. These topographical maps, or other constructed three-dimensional maps, may prove to be an interesting area of research with regards to generalizing design to multiple dimensions.

Finally, the idea of multidimensional hierarchical graphical interfaces could be explored with other physical subject areas. One such idea would be to use a multidimensional interface for exploring the human body, where each level of depth would show a more specific view of the human body from skin to muscle to the circulatory system to bones and so on.

4.3 Conclusions

The results highlight the influence of a multidimensional map-based interface on subject efficiency with respect to response time and accuracy. While the depth and breadth of the interface did not seem to matter, subjects performed better using the graphical representation interface without scent when faced with complex tasks. The results also showed that as the depth levels increase in a multidimensional map-based interface, the presence of labels aided navigation.

Based on the study completed, it is recommended that future constructs of multidimensional graphical map-based interfaces rely on the representation of the geography when designing broad interfaces. However when designing deep interfaces, constructs should include map labels to aid in navigation.

REFERENCES

Print Publications:

[Furnas97]
Furnas, G. W., "Effective view navigation," Proceedings of the SIGCHI conference on Human factors in computing systems 1, 1 (March 1997), pp. 367-374.

[Gray86]
Gray, J., "The role of menu titles as a navigational aid in hierarchical menus," SIGCHI Bulletin 17, 3 (January, 1986) pp. 33-40.

[Hornbaek02]
Hornbaek, K., Bederson, B. B., Plaisant, C., "Navigation patterns and usability of zoomable user interfaces with and without an overview," ACM Transactions on Computer-Human Interaction (TOCHI) 9, 4 (December 2002), pp. 362-389.

[Landauer85]
Landauer, T.K. and D.W. Nachbar, "Selection from alphabetic and numeric menu trees using a touch screen: breadth, depth, and widtb," Proceedings of CHI '85 Human Factors in Computing Systems 1, 1 (April, 1985), pp. 73-78.

[Larson98]
Larson, K. and M. Czerwinski, "Web page design: implications of memory, structure and scent for information retrieval," Proceedings of CHI '98 Human Factors in Computing Systems 1, 1 (April, 1998) pp. 25-32.

[Lim02]
Lim, E., Goh, D. H., Liu, Z., Ng, W., Khoo, C.S., Higgins, S. E., "G-Portal: A map-based digital library for distributed geospatial and georeferenced resources," Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries, 1, 1 (July, 2002), pp. 351-358.

[Miller56]
Miller, G.A., "The magical number seven, plus or minus two: some limits on our capacity for processing information," The Psychological Review 63, 1 (1956), pp. 81-97.

[Shneiderman05]
Shneiderman, B. and C. Plaisant, Designing the User Interface: Strategies for Effective Human-Computer Interaction, Addison-Wesley Publishing Co., Inc., Boston, 2005.

[Tullis85]
Tullis, T.S., "Designing a menu-based interface to an operating system," Proceedings of CHI '85 Human Factors in Computing Systems 1, 1 (April, 1985) pp. 79-84.

Electronic Sources:

[AdobeIllustrator05]
Adobe Illustrator CS2, 2005, download from http://www.adobe.com/downloads, last accessed December 1, 2006.

[AdobePhotoshop05]
Adobe Photoshop CS2, 2005, download from http://www.adobe.com/downloads, last accessed December 19, 2006.

[CIA06]
CIA The World Factbook, 2006, download from https://www.cia.gov/cia/publications/factbook/docs/refmaps.html, last accessed November 12, 2006.

[MacromediaDreamweaver05]
Macromedia Dreamweaver 8, 2005, download from http://www.adobe.com/downloads, last accessed March 1, 2007.

APPENDIX A

List of Simple Questions

What is the population of Barrow, Alaska, United States of America?

How high is the plateau on which the city of Brasilia, Brazil is built?

In what year was Durban, South Africa established?

What is the population of Magadan, Russia?

In what year was New Delhi, India converted into a Union territory?

APPENDIX B

List of Compound Questions

What is greater, the population of Baghdad, Iraq or the number of daily subway passengers in Seoul, South Korea?

What is longer, the Dnieper River that flows through Kiev, Ukraine, or the Tagus River that flows through Lisbon, Portugal?

Which is greater, the number of boroughs in London, United Kingdom or the number of communes in Dakar, Senegal?

What is greater, the population of Toronto, Canada or Auckland, New Zealand?

What is greater, the population of Manila, Philippines or Bogota, Colombia?

APPENDIX C

List of Demographic and Preference Questions

What is your gender?
() Female () Male

What is your age?
() Under 18 () 18-24 () 25-31 () 32-38 () 39-45 () 46-54 () Over 55

Are you right or left handed?
() Left () Right

Do you have 20/20 vision naturally or through corrected lenses?
() Yes () No

What is your level of computer experience?
() Beginner () Intermediate () Advanced

What is your level of education?
() Some High School () High School/GED () Some College () Associate's Degree () Bachelor's Degree () Graduate Degree

For the following questions, please answer on the corresponding scale of 1 to 5 where 1 is Strongly Disagree and 5 is Strongly Agree.

I did not feel lost finding the cities.
() 1 - Strongly Disagree () 2 - Disagree () 3 - Neither Agree Nor Disagree () 4 - Agree () 5 - Strongly Agree

I was never at a loss as to what the next step was.
() 1 - Strongly Disagree () 2 - Disagree () 3 - Neither Agree Nor Disagree () 4 - Agree () 5 - Strongly Agree

It was easy to find the information I was looking for.
() 1 - Strongly Disagree () 2 - Disagree () 3 - Neither Agree Nor Disagree () 4 - Agree () 5 - Strongly Agree

I liked using the interface.
() 1 - Strongly Disagree () 2 - Disagree () 3 - Neither Agree Nor Disagree () 4 - Agree () 5 - Strongly Agree

I felt confident using this interface.
() 1 - Strongly Disagree () 2 - Disagree () 3 - Neither Agree Nor Disagree () 4 - Agree () 5 - Strongly Agree

The interface was well organized.
() 1 - Strongly Disagree () 2 - Disagree () 3 - Neither Agree Nor Disagree () 4 - Agree () 5 - Strongly Agree

I would rather use this interface than a text-based interface.
() 1 - Strongly Disagree () 2 - Disagree () 3 - Neither Agree Nor Disagree () 4 - Agree () 5 - Strongly Agree

APPENDIX D

Subject Group Frequencies

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	F	58	65.2	65.2	65.2
	M	31	34.8	34.8	100.0
	Total	89	100.0	100.0	

Handedness

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	L	8	9.0	9.0	9.0
	R	81	91.0	91.0	100.0
	Total	89	100.0	100.0	

Eye 20/20

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	N	31	34.8	34.8	34.8
	Y	58	65.2	65.2	100.0
	Total	89	100.0	100.0	

Computer Experience

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Advanced	26	29.2	29.2	29.2
	Beginner	8	9.0	9.0	38.2
	Intermediate	55	61.8	61.8	100.0
	Total	89	100.0	100.0	

APPENDIX E

Subject Group Crosstabs

gender * int format * int label Crosstabulation
Count

int label			int format			Total
			4X4X4	8X8	16X4	4X4X4
WITH	gender	F	13	13	11	37
		M	2	4	2	8
	Total		15	17	13	45
WITHOUT	gender	F	5	9	7	21
		M	10	4	9	23
	Total		15	13	16	44

hand * int format * int label Crosstabulation
Count

int label			int format			Total
			4X4X4	8X8	16X4	4X4X4
WITH	hand	L	2	3	1	6
		R	13	14	12	39
	Total		15	17	13	45
WITHOUT	hand	L	1	1	0	2
		R	14	12	16	42
	Total		15	13	16	44

eye * int format * int label Crosstabulation
Count

int label			int format			Total
			4X4X4	8X8	16X4	4X4X4
WITH	eye	N	3	5	4	12
		Y	12	12	9	33
	Total		15	17	13	45
WITHOUT	eye	N	7	4	8	19
		Y	8	9	8	25
	Total		15	13	16	44

compexp * int format * int label Crosstabulation
Count

int label			int format			Total
			4X4X4	8X8	16X4	4X4X4
WITH	compexp	Advanced	2	5	1	8
		Beginner	1	1	2	4
		Intermediate	12	11	10	33
	Total		15	17	13	45
WITHOUT	compexp	Advanced	8	4	6	18
		Beginner	2	2	0	4
		Intermediate	5	7	10	22
	Total		15	13	16	44

APPENDIX F

Significant Correlations

Correlation Pair		r	p
q1 clicks simple	q2 clicks simple	0.635	0.000
q1 clicks simple	q2 errors simple	0.619	0.000
q1 clicks simple	q2 start to finish time simple	0.641	0.000
q1 clicks simple	q3 clicks simple	0.327	0.002
q1 clicks simple	q3 errors simple	0.324	0.002
q1 clicks simple	q4 start to finish time simple	0.318	0.002
q1 clicks simple	q5 clicks simple	0.347	0.001
q1 clicks simple	q5 errors simple	0.341	0.001
q1 clicks simple	q5 start to finish time simple	0.323	0.002
q1 clicks simple	p2 - I was never at a loss to what the next step was	-0.359	0.001
q1 clicks simple	q4 answer incorrect simple	-0.504	0.000
q1 clicks simple	q5 answer incorrect simple	-0.504	0.000
q1 clicks simple	total answers compound	-0.423	0.000
q1 clicks simple	q4 answer correctness compound	-0.315	0.003

Correlation Pair		r	p
q1 clicks simple	q5 answer correctness compound	-0.561	0.000
q1 errors simple	q1 start to finish time simple	0.578	0.000
q1 errors simple	q2 clicks simple	0.652	0.000
q1 errors simple	q2 errors simple	0.641	0.000
q1 errors simple	q2 start to finish time simple	0.662	0.000
q1 errors simple	q3 clicks simple	0.320	0.002
q1 errors simple	q3 errors simple	0.325	0.002
q1 errors simple	q4 start to finish time simple	0.345	0.001
q1 errors simple	q5 clicks simple	0.365	0.000
q1 errors simple	q5 errors simple	0.365	0.000
q1 errors simple	q5 start to finish time simple	0.343	0.001
q1 errors simple	p2 - I was never at a loss to what the next step was	-0.358	0.001
q1 errors simple	q4 answer incorrect simple	-0.501	0.000
q1 errors simple	q5 answer incorrect simple	-0.501	0.000
q1 errors simple	total answers compound	-0.426	0.000
q1 errors simple	q4 answer correctness compound	-0.309	0.003
q1 errors simple	q5 answer correctness compound	-0.562	0.000
q1 start to finish time simple	q2 clicks simple	0.518	0.000

Correlation Pair		r	p
q1 start to finish time simple	q2 errors simple	0.499	0.000
q1 start to finish time simple	q2 start to finish time simple	0.530	0.000
q1 start to finish time simple	q5 clicks simple	0.346	0.000
q1 start to finish time simple	q5 errors simple	0.337	0.001
q1 start to finish time simple	q5 start to finish time simple	0.312	0.003
q1 start to finish time simple	q1 first click time compound	0.613	0.000
q1 start to finish time simple	p2 - I was never at a loss to what the next step was	-0.360	0.001
q1 start to finish time simple	q4 answer incorrect simple	-0.359	0.001
q1 start to finish time simple	q5 answer incorrect simple	-0.359	0.001
q1 start to finish time simple	q5 answer correctness compound	-0.450	0.000
q1 first click time simple	q2 answer incorrect simple	-0.319	0.002
q2 clicks simple	q3 clicks simple	0.476	0.000
q2 clicks simple	q3 errors simple	0.485	0.000
q2 clicks simple	q3 start to finish time simple	0.405	0.000
q2 clicks simple	q4 start to finish time simple	0.505	0.000
q2 clicks simple	q5 clicks simple	0.569	0.000
q2 clicks simple	q5 errors simple	0.567	0.000
q2 clicks simple	q5 start to finish time simple	0.541	0.000

Correlation Pair		r	p
q2 clicks simple	q5 clicks compound	0.327	0.002
q2 clicks simple	q5 errors compound	0.328	0.002
q2 clicks simple	p2 - I was never at a loss to what the next step was	-0.471	0.000
q2 clicks simple	q1 answer incorrect simple	-0.580	0.000
q2 clicks simple	q4 answer incorrect simple	-0.580	0.000
q2 clicks simple	q5 answer incorrect simple	-0.580	0.000
q2 clicks simple	total answers compound	-0.466	0.000
q2 clicks simple	q1 answer correctness compound	-0.302	0.004
q2 clicks simple	q4 answer correctness compound	-0.317	0.002
q2 clicks simple	q5 answer correctness compound	-0.740	0.000
q2 errors simple	q3 clicks simple	0.456	0.000
q2 errors simple	q3 errors simple	0.470	0.000
q2 errors simple	q3 start to finish time simple	0.389	0.000
q2 errors simple	q4 start to finish time simple	0.511	0.000
q2 errors simple	q5 clicks simple	0.568	0.000
q2 errors simple	q5 errors simple	0.569	0.000
q2 errors simple	q5 start to finish time simple	0.541	0.000
q2 errors simple	q5 clicks compound	0.327	0.002

Correlation Pair		r	p
q2 errors simple	q5 errors compound	0.337	0.001
q2 errors simple	p2 - I was never at a loss to what the next step was	-0.458	0.000
q2 errors simple	q1 answer incorrect simple	-0.550	0.000
q2 errors simple	q4 answer incorrect simple	-0.550	0.000
q2 errors simple	q5 answer incorrect simple	-0.550	0.000
q2 errors simple	total answers compound	-0.449	0.000
q2 errors simple	q5 answer correctness compound	-0.718	0.000
q2 start to finish time simple	q3 clicks simple	0.492	0.000
q2 start to finish time simple	q3 errors simple	0.505	0.000
q2 start to finish time simple	q3 start to finish time simple	0.434	0.000
q2 start to finish time simple	q4 clicks simple	0.337	0.001
q2 start to finish time simple	q4 errors simple	0.322	0.002
q2 start to finish time simple	q4 start to finish time simple	0.548	0.000
q2 start to finish time simple	q5 clicks simple	0.557	0.000
q2 start to finish time simple	q5 errors simple	0.556	0.000
q2 start to finish time simple	q5 start to finish time simple	0.545	0.000
q2 start to finish time simple	q5 clicks compound	0.323	0.002
q2 start to finish time simple	q5 errors compound	0.329	0.002

Correlation Pair		r	p
q2 start to finish time simple	q5 start to finish time compound	0.352	0.001
q2 start to finish time simple	p2 - I was never at a loss to what the next step was	-0.491	0.000
q2 start to finish time simple	q1 answer incorrect simple	-0.613	0.000
q2 start to finish time simple	q4 answer incorrect simple	-0.613	0.000
q2 start to finish time simple	q5 answer incorrect simple	-0.613	0.000
q2 start to finish time simple	total answers compound	-0.490	0.000
q2 start to finish time simple	q1 answer correctness compound	-0.325	0.002
q2 start to finish time simple	q4 answer correctness compound	-0.328	0.002
q2 start to finish time simple	q5 answer correctness compound	-0.753	0.000
q3 clicks simple	p1 - I did not feel lost finding the cities	-0.399	0.000
q3 clicks simple	p3 - It was easy to find the information I was looking for	-0.345	0.001
q3 clicks simple	p6 - The interface was well organized	-0.315	0.003
q3 clicks simple	q1 answer incorrect simple	-0.700	0.000
q3 clicks simple	q4 answer incorrect simple	-0.700	0.000
q3 clicks simple	q5 answer incorrect simple	-0.700	0.000
q3 clicks simple	total answers compound	-0.427	0.000
q3 clicks simple	q1 answer correctness compound	-0.376	0.000
q3 clicks simple	q4 answer correctness compound	-0.366	0.000

Correlation Pair		r	p
q3 clicks simple	q5 answer correctness compound	-0.443	0.000
q3 errors simple	p3 - It was easy to find the information I was looking for	-0.324	0.002
q3 errors simple	q1 answer incorrect simple	-0.679	0.000
q3 errors simple	q4 answer incorrect simple	-0.679	0.000
q3 errors simple	q5 answer incorrect simple	-0.679	0.000
q3 errors simple	total answers compound	-0.435	0.000
q3 errors simple	q1 answer correctness compound	-0.372	0.000
q3 errors simple	q4 answer correctness compound	-0.348	0.001
q3 errors simple	q5 answer correctness compound	-0.442	0.000
q3 start to finish time simple	p1 - I did not feel lost finding the cities	-0.361	0.001
q3 start to finish time simple	q1 answer incorrect simple	-0.568	0.000
q3 start to finish time simple	q4 answer incorrect simple	-0.568	0.000
q3 start to finish time simple	q5 answer incorrect simple	-0.568	0.000
q3 start to finish time simple	total answers compound	-0.350	0.001
q3 start to finish time simple	q1 answer correctness compound	-0.318	0.002
q3 start to finish time simple	q5 answer correctness compound	-0.380	0.000
q3 first click time simple	q5 first click time simple	0.428	0.000
q4 clicks simple	q5 clicks simple	0.435	0.000

Correlation Pair		r	p
q4 clicks simple	q5 errors simple	0.439	0.000
q4 clicks simple	q5 start to finish time simple	0.431	0.000
q4 clicks simple	q1 start to finish time compound	0.507	0.000
q4 clicks simple	q4 clicks compound	0.358	0.001
q4 clicks simple	q4 errors compound	0.347	0.001
q4 clicks simple	q4 start to finish time compound	0.303	0.001
q4 clicks simple	q5 clicks compound	0.380	0.000
q4 clicks simple	q5 errors compound	0.393	0.000
q4 clicks simple	q5 start to finish time compound	0.424	0.000
q4 clicks simple	p4 - I liked using the interface	-0.332	0.001
q4 errors simple	q5 clicks simple	0.378	0.000
q4 errors simple	q5 errors simple	0.387	0.000
q4 errors simple	q5 start to finish time simple	0.386	0.000
q4 errors simple	q1 start to finish time compound	0.510	0.000
q4 errors simple	q4 clicks compound	0.332	0.001
q4 errors simple	q4 errors compound	0.331	0.002
q4 errors simple	q4 start to finish time compound	0.300	0.001
q4 errors simple	q5 clicks compound	0.326	0.002

Correlation Pair		r	p
q4 errors simple	q5 errors compound	0.351	0.001
q4 errors simple	q5 start to finish time compound	0.397	0.000
q4 errors simple	p4 - I liked using the interface	-0.359	0.001
q4 errors simple	p5 - I felt confident using this interface	-0.330	0.002
q4 start to finish time simple	q5 clicks simple	0.646	0.000
q4 start to finish time simple	q5 errors simple	0.653	0.000
q4 start to finish time simple	q5 start to finish time simple	0.660	0.000
q4 start to finish time simple	q1 clicks compound	0.325	0.002
q4 start to finish time simple	q1 errors compound	0.350	0.001
q4 start to finish time simple	q2 first click time compound	0.312	0.003
q4 start to finish time simple	q4 clicks compound	0.403	0.000
q4 start to finish time simple	q4 errors compound	0.407	0.000
q4 start to finish time simple	q4 start to finish time compound	0.358	0.001
q4 start to finish time simple	q5 clicks compound	0.451	0.000
q4 start to finish time simple	q5 errors compound	0.479	0.000
q4 start to finish time simple	q5 start to finish time compound	0.478	0.000
q4 start to finish time simple	q5 answer correctness compound	-0.484	0.000
q5 clicks simple	q1 clicks compound	0.447	0.000

Correlation Pair		r	p
q5 clicks simple	q1 errors compound	0.475	0.000
q5 clicks simple	q3 clicks compound	0.407	0.000
q5 clicks simple	q3 errors compound	0.353	0.001
q5 clicks simple	q4 clicks compound	0.404	0.000
q5 clicks simple	q4 errors compound	0.396	0.000
q5 clicks simple	q4 start to finish time compound	0.312	0.003
q5 clicks simple	q5 clicks compound	0.435	0.000
q5 clicks simple	q5 errors compound	0.443	0.000
q5 clicks simple	q5 start to finish time compound	0.441	0.000
q5 clicks simple	p2 - I was never at a loss to what the next step was	-0.360	0.001
q5 clicks simple	q3 answer correctness compound	-0.302	0.004
q5 clicks simple	q5 answer incorrect complex	-0.477	0
q5 errors simple	q1 clicks compound	0.447	0.000
q5 errors simple	q1 errors compound	0.480	0.000
q5 errors simple	q3 clicks compound	0.398	0.000
q5 errors simple	q3 errors compound	0.344	0.001
q5 errors simple	q4 clicks compound	0.407	0.000
q5 errors simple	q4 errors compound	0.406	0.000

Correlation Pair		r	p
q5 errors simple	q4 start to finish time compound	0.320	0.002
q5 errors simple	q5 clicks compound	0.438	0.000
q5 errors simple	q5 errors compound	0.455	0.000
q5 errors simple	q5 start to finish time compound	0.448	0.000
q5 errors simple	p2 - I was never at a loss to what the next step was	-0.352	0.001
q5 errors simple	q5 answer correctness compound	-0.471	0.000
q5 start to finish time simple	q1 clicks compound	0.413	0.000
q5 start to finish time simple	q1 errors compound	0.439	0.000
q5 start to finish time simple	q3 clicks compound	0.378	0.000
q5 start to finish time simple	q3 errors compound	0.324	0.002
q5 start to finish time simple	q4 clicks compound	0.408	0.000
q5 start to finish time simple	q4 errors compound	0.405	0.000
q5 start to finish time simple	q4 start to finish time compound	0.373	0.000
q5 start to finish time simple	q4 first click time compound	0.307	0.003
q5 start to finish time simple	q5 clicks compound	0.399	0.000
q5 start to finish time simple	q5 errors compound	0.413	0.000
q5 start to finish time simple	q5 start to finish time compound	0.476	0.000
q5 start to finish time simple	p2 - I was never at a loss to what the next step was	-0.331	0.002

Correlation Pair		r	p
q5 start to finish time simple	q5 answer correctness compound	-0.476	0.000
q5 first click time simple	q3 first click time compound	0.323	0.002
q1 clicks compound	q2 clicks compound	0.487	0.000
q1 clicks compound	q2 errors compound	0.509	0.000
q1 clicks compound	q3 clicks compound	0.663	0.000
q1 clicks compound	q3 errors compound	0.611	0.000
q1 clicks compound	q3 start to finish time compound	0.474	0.000
q1 clicks compound	q4 clicks compound	0.335	0.001
q1 clicks compound	q4 errors compound	0.338	0.001
q1 clicks compound	p1 - I did not feel lost finding the cities	-0.407	0.000
q1 errors compound	q2 clicks compound	0.468	0.000
q1 errors compound	q2 errors compound	0.499	0.000
q1 errors compound	q3 clicks compound	0.666	0.000
q1 errors compound	q3 errors compound	0.610	0.000
q1 errors compound	q3 start to finish time compound	0.471	0.000
q1 errors compound	q4 clicks compound	0.319	0.002
q1 errors compound	q4 errors compound	0.333	0.001
q1 errors compound	p1 - I did not feel lost finding the cities	-0.380	0.000

Correlation Pair		r	p
q2 clicks compound	q3 clicks compound	0.337	0.001
q2 clicks compound	q3 errors compound	0.325	0.002
q2 clicks compound	q3 start to finish time compound	0.357	0.010
q2 clicks compound	q4 clicks compound	0.514	0.000
q2 clicks compound	q4 errors compound	0.523	0.000
q2 clicks compound	q4 start to finish time compound	0.432	0.000
q2 clicks compound	q5 clicks compound	0.537	0.000
q2 clicks compound	q5 errors compound	0.570	0.000
q2 clicks compound	q5 start to finish time compound	0.323	0.002
q2 errors compound	q3 clicks compound	0.362	0.000
q2 errors compound	q3 errors compound	0.350	0.001
q2 errors compound	q3 start to finish time compound	0.399	0.000
q2 errors compound	q4 clicks compound	0.508	0.000
q2 errors compound	q4 errors compound	0.527	0.000
q2 errors compound	q4 start to finish time compound	0.461	0.000
q2 errors compound	q5 clicks compound	0.493	0.000
q2 errors compound	q5 errors compound	0.540	0.000
q2 errors compound	q5 start to finish time compound	0.331	0.002

Correlation Pair		r	p
q2 start to finish time compound	q3 start to finish time compound	0.425	0.000
q2 start to finish time compound	q4 clicks compound	0.390	0.000
q2 start to finish time compound	q4 errors compound	0.411	0.000
q2 start to finish time compound	q4 start to finish time compound	0.511	0.000
q2 start to finish time compound	q5 errors compound	0.322	0.002
q2 start to finish time compound	q5 start to finish time compound	0.370	0.000
q2 start to finish time compound	p4 - I liked using the interface	-0.380	0.000
q2 start to finish time compound	p5 - I felt confident using this interface	-0.302	0.004
q2 start to finish time compound	q1 answer correctness compound	-0.317	0.002
q3 clicks compound	q4 clicks compound	0.412	0.000
q3 clicks compound	q4 errors compound	0.412	0.000
q3 clicks compound	q4 start to finish time compound	0.376	0.000
q3 clicks compound	p1 - I did not feel lost finding the cities	-0.446	0.000
q3 clicks compound	p3 - It was easy to find the information I was looking for	-0.459	0.000
q3 clicks compound	p5 - I felt confident using this interface	-0.352	0.001
q3 errors compound	q4 clicks compound	0.392	0.000
q3 errors compound	q4 errors compound	0.393	0.000
q3 errors compound	q4 start to finish time compound	0.347	0.001

Correlation Pair		r	p
q3 errors compound	p1 - I did not feel lost finding the cities	-0.474	0.000
q3 errors compound	p3 - It was easy to find the information I was looking for	-0.475	0.000
q3 errors compound	p5 - I felt confident using this interface	-0.386	0.000
q3 start to finish time compound	q4 clicks compound	0.401	0.000
q3 start to finish time compound	q4 errors compound	0.414	0.000
q3 start to finish time compound	q4 start to finish time compound	0.492	0.000
q3 start to finish time compound	p1 - I did not feel lost finding the cities	-0.305	0.004
q3 start to finish time compound	p3 - It was easy to find the information I was looking for	-0.412	0.000
q4 clicks compound	q5 clicks compound	0.677	0.000
q4 clicks compound	q5 errors compound	0.677	0.000
q4 clicks compound	q5 start to finish time compound	0.469	0.000
q4 clicks compound	p2 - I was never at a loss to what the next step was	-0.368	0.000
q4 errors compound	q5 clicks compound	0.650	0.000
q4 errors compound	q5 errors compound	0.667	0.000
q4 errors compound	q5 start to finish time compound	0.456	0.000
q4 errors compound	p2 - I was never at a loss to what the next step was	-0.364	0.000
q4 start to finish time compound	q5 clicks compound	0.398	0.000
q4 start to finish time compound	q5 errors compound	0.407	0.000

Correlation Pair		r	p
q4 start to finish time compound	q5 start to finish time compound	0.442	0.000
q4 start to finish time compound	p2 - I was never at a loss to what the next step was	-0.340	0.001
q4 start to finish time compound	p4 - I liked using the interface	-0.353	0.001
q4 first click time compound	q5 start to finish time compound	0.582	0.000
q4 first click time compound	q5 first click time compound	0.788	0.000
q4 first click time compound	q5 answer correctness compound	-0.305	0.004
p1 - I did not feel lost finding the cities	p3 - It was easy to find the information I was looking for	0.465	0.000
p1 - I did not feel lost finding the cities	p5 - I felt confident using this interface	0.393	0.000
p1 - I did not feel lost finding the cities	total clicks simple	-0.309	0.000
p1 - I did not feel lost finding the cities	total clicks compound	-0.391	0.000
p1 - I did not feel lost finding the cities	total errors compound	-0.392	0.000
p1 - I did not feel lost finding the cities	total answers simple	0.303	0.004
p1 - I did not feel lost finding the cities	total clicks subject made	-0.427	0.000
p1 - I did not feel lost finding the cities	total errors subject made	-0.419	0.000
p1 - I did not feel lost finding the cities	total correct answers subject had	0.308	0.003
p1 - I did not feel lost finding the cities	total time taken by subject	-0.319	0.002
p2 - I was never at a loss to what the next step was	p3 - It was easy to find the information I was looking for	0.403	0.000
p2 - I was never at a loss to what the next step was	p5 - I felt confident using this interface	0.432	0.000

Correlation Pair		r	p
p2 - I was never at a loss to what the next step was	p6 - The interface was well organized	0.427	0.000
p2 - I was never at a loss to what the next step was	total clicks simple	-0.470	0.000
p2 - I was never at a loss to what the next step was	total errors simple	-0.462	0.000
p2 - I was never at a loss to what the next step was	total clicks compound	-0.331	0.002
p2 - I was never at a loss to what the next step was	total errors compound	-0.339	0.001
p2 - I was never at a loss to what the next step was	q1 answer incorrect simple	0.384	0.000
p2 - I was never at a loss to what the next step was	q4 answer incorrect simple	0.384	0.000
p2 - I was never at a loss to what the next step was	q5 answer incorrect simple	0.384	0.000
p2 - I was never at a loss to what the next step was	total answers compound	0.497	0.000
p2 - I was never at a loss to what the next step was	q1 answer correctness compound	0.324	0.002
p2 - I was never at a loss to what the next step was	q5 answer correctness compound	0.533	0.000
p2 - I was never at a loss to what the next step was	total answers simple	0.449	0.000
p2 - I was never at a loss to what the next step was	total time simple	-0.481	0.000
p2 - I was never at a loss to what the next step was	total clicks subject made	-0.448	0.000
p2 - I was never at a loss to what the next step was	total errors subject made	-0.450	0.000
p2 - I was never at a loss to what the next step was	total correct answers subject had	0.530	0.000
p3 - It was easy to find the information I was looking for	p4 - I liked using the interface	0.482	0.000
p3 - It was easy to find the information I was looking for	p5 - I felt confident using this interface	0.639	0.000

Correlation Pair		r	p
p3 - It was easy to find the information I was looking for	p6 - The interface was well organized	0.454	0.000
p3 - It was easy to find the information I was looking for	p7 - I would rather use this interface than a text-based interface	0.575	0.000
p3 - It was easy to find the information I was looking for	total clicks compound	-0.384	0.000
p3 - It was easy to find the information I was looking for	total errors compound	-0.396	0.000
p3 - It was easy to find the information I was looking for	q1 answer correctness compound	0.351	0.001
p3 - It was easy to find the information I was looking for	total time compound	-0.387	0.000
p3 - It was easy to find the information I was looking for	total clicks subject made	-0.398	0.000
p3 - It was easy to find the information I was looking for	total errors subject made	-0.403	0.000
p3 - It was easy to find the information I was looking for	total time taken by subject	-0.413	0.000
p4 - I liked using the interface	p5 - I felt confident using this interface	0.640	0.000
p4 - I liked using the interface	p6 - The interface was well organized	0.585	0.000
p4 - I liked using the interface	p7 - I would rather use this interface than a text-based interface	0.472	0.000
p4 - I liked using the interface	q1 answer correctness compound	0.312	0.003
p4 - I liked using the interface	total time compound	-0.419	0.000
p4 - I liked using the interface	total time taken by subject	-0.407	0.000
p5 - I felt confident using this interface	p6 - The interface was well organized	0.624	0.000
p5 - I felt confident using this interface	p7 - I would rather use this interface than a text-based interface	0.413	0.000

Correlation Pair		r	p
p5 - I felt confident using this interface	total errors compound	-0.303	0.004
p5 - I felt confident using this interface	q1 answer correctness compound	0.399	0.000
p5 - I felt confident using this interface	total time compound	-0.370	0.000
p5 - I felt confident using this interface	total clicks subject made	-0.303	0.004
p5 - I felt confident using this interface	total errors subject made	-0.307	0.003
p5 - I felt confident using this interface	total time taken by subject	-0.367	0.000
p6 - The interface was well organized	p7 - I would rather use this interface than a text-based interface	0.471	0.000
p6 - The interface was well organized	q1 answer incorrect simple	0.301	0.004
p6 - The interface was well organized	q4 answer incorrect simple	0.301	0.004
p6 - The interface was well organized	a5asn	0.301	0.004
p6 - The interface was well organized	total answers compound	0.303	0.004
p6 - The interface was well organized	q1 answer correctness compound	0.366	0.000
p6 - The interface was well organized	total correct answers subject had	0.332	0.001
p7 - I would rather use this interface than a text-based interface	q1 answer correctness compound	0.309	0.003
total clicks simple	q1 answer incorrect simple	-0.452	0.000
total clicks simple	q4 answer incorrect simple	-0.452	0.000
total clicks simple	q5 answer incorrect simple	-0.452	0.000

Correlation Pair		r	p
total clicks simple	total answers compound	-0.434	0.000
total clicks simple	q5 answer correctness compound	-0.688	0.000
total clicks simple	total answers simple	-0.420	0.000
total clicks simple	total correct answers subject had	-0.475	0.000
total errors simple	q1 answer incorrect simple	-0.431	0.000
total errors simple	q4 answer incorrect simple	-0.431	0.000
total errors simple	q5 answer incorrect simple	-0.431	0.000
total errors simple	total answers compound	-0.431	0.000
total errors simple	q5 answer correctness compound	-0.671	0.000
total errors simple	total answers simple	-0.403	0.000
total errors simple	total correct answers subject had	-0.465	0.000
total clicks compound	q2 answer correctness compound	-0.306	0.004
q2 answer incorrect simple	q2 answer correctness compound	0.339	0.001
q2 answer incorrect simple	total answers simple	0.306	0.003
q1 answer incorrect simple	q4 answer incorrect simple	1.000	0.000
q1 answer incorrect simple	q5 answer incorrect simple	1.000	0.000
q1 answer incorrect simple	total answers compound	0.585	0.000
q1 answer incorrect simple	q1 answer correctness compound	0.571	0.000

Correlation Pair		r	p
q1 answer incorrect simple	q4 answer correctness compound	0.571	0.000
q1 answer incorrect simple	q5 answer correctness compound	0.571	0.000
q1 answer incorrect simple	total answers simple	0.948	0.000
q1 answer incorrect simple	total time simple	-0.430	0.000
q1 answer incorrect simple	total correct answers subject had	0.808	0.000
q4 answer incorrect simple	q5 answer incorrect simple	1.000	0.000
q4 answer incorrect simple	total answers compound	0.585	0.000
q4 answer incorrect simple	q1 answer correctness compound	0.571	0.000
q4 answer incorrect simple	q4 answer correctness compound	0.571	0.000
q4 answer incorrect simple	q5 answer correctness compound	0.571	0.000
q4 answer incorrect simple	total answers simple	0.948	0.000
q4 answer incorrect simple	total time simple	-0.430	0.000
q4 answer incorrect simple	total correct answers subject had	0.808	0.000
q5 answer incorrect simple	total answers compound	0.585	0.000
q5 answer incorrect simple	q1 answer correctness compound	0.571	0.000
q5 answer incorrect simple	q4 answer correctness compound	0.571	0.000
q5 answer incorrect simple	q5 answer correctness compound	0.571	0.000
q5 answer incorrect simple	total answers simple	0.948	0.000

Correlation Pair		r	p
q5 answer incorrect simple	total time simple	-0.430	0.000
q5 answer incorrect simple	total correct answers subject had	0.808	0.000
total answers compound	q1 answer correctness compound	0.535	0.000
total answers compound	q2 answer correctness compound	0.564	0.000
total answers compound	q4 answer correctness compound	0.535	0.000
total answers compound	q5 answer correctness compound	0.657	0.000
total answers compound	total answers simple	0.609	0.000
total answers compound	total time simple	-0.428	0.000
total answers compound	total clicks subject made	-0.387	0.000
total answers compound	total errors subject made	-0.378	0.000
total answers compound	total correct answers subject had	0.938	0.000
q1 answer correctness compound	q4 answer correctness compound	0.310	0.003
q1 answer correctness compound	q5 answer correctness compound	0.310	0.003
q1 answer correctness compound	total answers simple	0.537	0.000
q1 answer correctness compound	total correct answers subject had	0.594	0.000
q2 answer correctness compound	total correct answers subject had	0.411	0.000
q4 answer correctness compound	q5 answer correctness compound	0.310	0.003
q4 answer correctness compound	total answers simple	0.537	0.000

Correlation Pair		r	p
q4 answer correctness compound	total correct answers subject had	0.594	0.000
q5 answer correctness compound	total answers simple	0.537	0.000
q5 answer correctness compound	total time simple	-0.722	0.000
q5 answer correctness compound	total clicks subject made	-0.383	0.000
q5 answer correctness compound	total errors subject made	-0.365	0.000
q5 answer correctness compound	total correct answers subject had	0.676	0.000
total answers simple	total time simple	-0.422	0.000
total answers simple	total correct answers subject had	0.847	0.000
total time simple	total clicks subject made	0.656	0.000
total time simple	total errors subject made	0.650	0.000
total time simple	total correct answers subject had	-0.472	0.000
total time simple	total time taken by subject	0.431	0.000
total time compound	total clicks subject made	0.558	0.000
total time compound	total errors subject made	0.553	0.000
total time compound	total time taken by subject	0.969	0.000
total clicks subject made	total correct answers subject had	-0.331	0.002
total errors subject made	total correct answers subject had	-0.321	0.002

APPENDIX G

Means for ANOVA Results in Label and Non-Label Interfaces for Simple and Compound Questions

		Simple	Compound
With Labels			
	Total Errors	4.89	12.45
	Total Time	97.41	457.30
	Total Clicks	32.50	79.89
Without Labels			
	Total Errors	6.34	8.03
	Total Time	114.63	281.89
	Total Clicks	35.53	71.42

APPENDIX H

Means for ANOVA Results in Simple and Compound Questions

	Simple	Compound
Total Errors	5.56	10.40
Total Time	105.39	376.01
Total Clicks	33.90	75.96
Total Correct Answers	4.99	4.85

APPENDIX I

Means for ANOVA Results in Label and Non-Label Interfaces

	With Labels	Without Labels
Total Time	277.35	198.26

APPENDIX J

Means for ANOVA Results in Label and Non-Label Interfaces with Varying Depths and Breadths for Simple and Compound Questions

		With Labels	Without Labels
4x4x4	Simple	5.00	5.00
	Compound	5.00	4.64
8x8	Simple	4.94	5.00
	Compound	4.88	4.91
16x4	Simple	5.00	5.00
	Compound	4.77	4.92

VITA

Brian Malek received a Bachelor of Science degree from the University of North Florida (UNF) in Computer and Information Sciences in December, 2004, and expects to receive a Master of Science in Computer and Information Sciences, Software Engineering Track, from UNF in May, 2007. Dr. F. Layne Wallace is serving as his thesis advisor. Brian is currently employed as a Management Trainee at CSX Technology and has been with the company for almost two years. In January 2007, Brian was awarded the honor of Outstanding Alumnus at UNF for his work in designing a software system for Volunteers in Medicine, a local non-profit organization that provides primary medical services to the working uninsured of the Greater Jacksonville area.

Brian has on-going interests in human computer interfaces and adaptive hypermedia and has extensive programming experience in Java and .NET. His academic work has included the use of C, Java, COBOL, .NET, and SQL. Other interests include graphic design and digital photography. Married for almost two years, Brian spends his free time with his wife Jamie and their two dogs Pebbles and Bam Bam.