IOWA STATE UNIVERSITY
Digital Repository

Graduate Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

2008

Relationships between user performance and spatial ability in using map-based software on pen-based devices

Michelle Lenae Rusch
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Computer Sciences Commons

Relationships between user performance and spatial ability

in using map-based software on pen-based devices

by

Michelle Lenae Rusch

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Human Computer Interaction

Program of Study Committee:
Leslie Miller, Co-major Professor
Sarah Nusser, Co-major Professor
James Bernard
Reginald Souleyrette
Judy Vance

Iowa State University

Ames, Iowa

2008

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This thesis explored the association between individual differences and software user performance. The primary goal was to investigate the relationship between spatial ability and user performance in the context of map software. Two studies were conducted to investigate this goal. There were two hypotheses that applied to both user studies. First, we expected that performance measures such as time to complete a task or accuracy would be positively correlated with spatial ability, as measured by scores on cognitive assessments for sub-factors of spatial ability. Second, we hypothesized that differences between subjects with low spatial ability and subjects with high spatial ability would be more pronounced with complex tasks and interfaces, such as a map interface or a task to compare ground settings with a map, relative to simpler tasks, such as finding a button that matched a target color or using map software with guidance on how to use the software to execute the task protocol. Each user study also incorporated additional questions unique to the study. Our first user study investigated how the relationship between spatial ability and performance might be affected by the relevance of the button label and alternative button layouts. This research indicated that good button labeling helps to mitigate performance differences among subjects with varying spatial ability. Additionally, we found support that performance differences related to spatial ability were most apparent in complex conditions. The second study investigated whether the association between spatial ability and user performance was affected by offering guidance on performing a task and whether orientation and visualization were dissociable factors for a task involving ground to map comparisons. While there were no benefits found from the guidance, visualization and orientation skills were found to be

dissociable from one another for accuracy and tool use. Findings from both studies

underscore the importance of interface design in potentially reducing costs and burden for

users of map-based software in a small screen computing environment.

CHAPTER ONE. INTRODUCTION

In recent years, the field of human-computer interaction (HCI) has increasingly focused on the role of individual differences related to the cognitive abilities in software design. Benyon, Crerar, and Wilkinson (2001) suggested that differences in cognitive abilities are important because users with different levels of abilities tend to vary in their interpretation and use of software interfaces. Their perspective is that a software interface is essentially an arrangement of symbols presented to the end user. Because the user is restricted from seeing the underlying background of the system, the onus is placed on them to interpret machine behavior associated with interface symbols. Cognitive abilities influence a user's capacity and approach to developing strategies for interpreting the interface and its functions. For this reason, Benyon et al. (2001) stress the importance of considering these individual differences in the area of HCI.

Individual differences range from physiological (e.g., hand dominance) to psychological (e.g., memory). Researchers in HCI have explored a variety of individual differences that affect software user performance, including spatial ability, memory, reasoning abilities, verbal aptitude, and personality differences (Egan, 1988). While many of these differences are relevant, researchers have found spatial ability to be one of the strongest predictors of human computer performance (Dahlbäck, Höök, & Sjölinder, 1996; Egan, 1988; Egan & Gomez, 1985; Vicente, Hayes, & Williges, 1987; Pak, Rogers & Fisk, 2006; Büring, Gerken, & Reiterer, 2006). Lohman (1996) defines spatial ability as "the ability to generate, retain, retrieve, and transform well-structured visual images" (p. 98). The correlation between spatial ability and user performance is thought to occur because many

software functionalities rely on spatial reasoning, such as navigation via hypertext links or drop-down menus. Thus, software designs that recognize and accommodate variation in spatial abilities have the potential to improve user performance.

Limited research has been conducted to evaluate the relationship between spatial ability and user performance in the context of map interfaces. The use of digital maps involves spatial cognition because it involves the use of content that is explicitly geographic. An individual's spatial ability has an impact on how s/he acquires geographic knowledge and conceptualizes geographic space (Devlin & Berstein, 1995; Golledge, 1991; Mark, Freska, Hirtle, Lloyd, & Tversky, 1999; Siegel & White, 1975). For example, persons with better spatial visualization ability are better able to interpret map content and develop a mental configuration to efficiently navigate from one point to another. In addition, persons with better spatial orientation are most effective at completing tasks that involve comparing (paper or digital) map content to the surrounding environment.

Research by Büring et al. (2006) suggested that spatial ability is related to user preferences for scatterplot software interface designs. Their study compared performance for zoomable interfaces that displayed an overview window with similar interfaces that did not include an overview window. They found that persons with low spatial ability showed better performance when the overview was present relative to when the interface did not include the overview window. In contrast, high spatial ability users exhibited poorer performance for the interface that included the overview window relative to the interface without the overview window. This research suggests that spatial ability may influence whether aids or guidance provided by an interface are helpful or present a distraction.

To explore a relationship between user performance and spatial ability in the use of map-based software, we conducted two studies to investigate user responses to software interfaces involving maps. The first study examined spatial ability and user performance in the context of learning the locations of buttons associated with software functions. There were two treatments. The label treatments included four different kinds of labels that varied from highly informative to uninformative. The interface layout treatments included three interfaces that displayed buttons that executed a simple color-matching task in a circular arrangement, a grid layout of the buttons for the same task, and a grid layout of buttons associated with map functions that changed the map displayed in the interface above the buttons.

The second study investigated spatial ability in relation to user performance for a map interface that offered protocol guidance in performing a complex task in comparison with a similar interface that offered no guidance. The task required users to follow a sequential procedure to make a decision about whether the location of a housing unit was accurately represented by a digital map. The task was based on an address canvassing procedure used by the United States Census Bureau to prepare address lists and maps for the decennial census. The address canvassing task was especially useful for investigating two critical sub-factors of spatial ability: spatial visualization and spatial orientation. Both of these abilities are needed in evaluating whether a map accurately reflects the location of a housing unit. Specifically, because matching a map to the real world involves multiple serial operations inclusive of mental rotations and storing visual memory. Recent research by Pak et al. (2006) indicates that visualization and orientation have different effects on performance, depending on the nature of the task. They found that spatial orientation ability was related to

performance with tasks that were high in their navigation requirement. While this group made progress in testing orientation and visualization, their assessment of the mental rotation aspect of orientation did not involve evaluation of the person's frame of reference (i.e., their understanding of where they were at in relationship to their physical surroundings).

For both studies, performance metrics included time to perform tasks and measures of accuracy (button location recall or placement of housing unit locations), as well as user behaviors (e.g., number of times a tool was used). We hypothesized that subjects with low levels of spatial ability would have poorer performance than subjects with high spatial ability. In addition, we hypothesized that associations between performance and spatial ability would be smaller for treatment conditions that were easier for subjects (e.g., the color match task in the first study, and the guided software condition in the second study). In the second study, we also hypothesized that visualization and orientation would be dissociable for some performance metrics.

1.1 Scientific contributions

This research resulted in the following contributions to the understanding of the correlation between spatial ability and user performance for map-based software interfaces.

1) We found evidence that spatial ability was related to software user performance for map-based interfaces, extending prior research that focused on other interface settings.

2) We also found that for tasks that involve map reading and perspective-taking activities, visualization and orientation were important as differentiable components of spatial ability.

3) We found evidence that differences in performance between subjects with low and high spatial ability depended on the simplicity of the setting or task. That is, performance differences tended to be larger in more complex settings.

CHAPTER 2. REVIEW OF LITERATURE

2.1 Introduction

This chapter presents an overview of relevant literature specific to the research introduced in Chapter 1. The following review is organized into three major sections that present relevant research in each of these areas. The first section looks at the general concepts of individual differences. The second section addresses the specifics of spatial ability. The third section presents relevant aspects of cognition related to spatial ability. The final section looks at guided and unguided user interfaces.

2.2 Individual differences

The term individual differences has been described as the physical, cognitive, psychological, social, or cultural background that defines a person's aptitudes, preferences, abilities, and attitudes (Benyon, Crerar, & Wilkinsin, 2001). Benyon et al. (2001) suggested that such differences are important to consider when designing software because of the implied strategy involved for interpreting technological devices. They look at HCI as a field that typically concerns the expression of abstract symbols that represent something else. Thus because the end user is presented with a presentation of symbols and most often are restricted from seeing the underlying background of the system they are left to their own devices (e.g., cognitive abilities) to interpret machine behavior. They noted that cognitive abilities influence a person's capacity to develop the essential strategies necessary for this interpretation.

Benyon et al. (2001) presented three categories for individual differences that include physiological, socio-cultural, and psychological characteristics. Physiological differences involve biological classifications such as gender, size, vision, hearing, and mobility. Socio-cultural differences involve characteristics related to language, cultural, and environmental aspects of an individual's surroundings. The limitation on investigating physiological and socio-cultural characteristics for software design is that they involve measures that are typically too broad to be useful. Understanding what distinguishes specific user groups in cognitive terms (i.e., psychological distinctions) can help to constrain the number of potential design solutions (Dillon & Watson, 1996).

Psychological distinctions can be broken into three broad categories: intelligence, cognitive style, and personality. Intelligence involves mental abilities (e.g., spatial ability, perceptual speed, and memory to name a few). Cognitive style represents stable attitudes, preferences, or habitual strategies that influence perceiving, remembering, thinking, and problem solving (Messick, 1976). A broad range of dimensions of cognitive styles has been proposed ranging from reflective-impulsive to visualize-verbalizer. Personality has been loosely defined as "traits or stable tendencies to respond to certain classes of stimuli or situations in predictable ways" (Dillon & Watson, 1996, p. 626). The thought patterns proposed for cognitive styles and personality remain hard to distinguish from one another (Dillon & Watson, 1996). Additionally, cognitive styles have been shown to have little value in predicting user performance on user interfaces (Booth, Fowler, & Macaulay, 1987). Dillon and Watson report that personality differences thus far have not provided much support in predicting user performance on user interfaces either. Individual differences related to intelligence however are turning up better results (Dillon & Watson, 1996).

An important issue in the context of individual differences is the question of changeability. Some characteristics, such as speech, personal knowledge, and behavior are subject to change over time. Other characteristics, like gender, eye color, and intelligence tend to be fixed. Characteristics that are more resistant to change are more useful as a means of categorizing users. Van der Veer, Tauber, Waern, and Van Muylwijk (1985) suggested that intelligence is one characteristic most resistant to change.

Age can have an impact on the question of changeability. As people age, it is not uncommon to see changes in the cognitive abilities associated with intelligence. Haigh (1993) stated that tasks that involve skills used on a regular basis for a long period of time (e.g., language) can remain intact well into the late stages of life. In contrast, skills such as logical or mathematical reasoning and spatial ability are more vulnerable to degradation because they might not get used as frequently.

Kalyuga, Chandler, and Sweller (1998) illustrated how individual differences can affect user performance when expertise is taken into account. They found that as expertise changed so did the type of user interface that was most efficient. Their main finding was that novice users preferred graphics (i.e., maps and diagrams) to be integrated with text that explained the graphics, whereas experts were more efficient with graphical diagrams that were not integrated with the extra text. Kalyuga et al.'s theory was that as a person's expertise changes so do their needs from the interface. They described the mental process of integration to be cognitively demanding. They proposed that adding extra information to the stimulus can be helpful to novice users understanding the material. In contrast, they discovered the reverse findings for expert users. This group experienced unnecessary cognitive load in processing the extra information because it was found to be redundant.

These findings are not surprising, since expert users would have more experience. This experience would allow the expert users to develop more advanced strategies than the novice users.

Egan (1988) found that user performance in accessing information via software interfaces is affected by an individual's spatial ability, memory, reasoning ability, verbal aptitude, and possibly personality. While Dillon and Watson (1996) reported that personality differences have shown little support in predicting performance on user interfaces, the other abilities suggested by Egan offer potential. Of these abilities, spatial ability has been found to be one of the strongest predictors of human computer performance (Dahlbäck, Höök, & Sjölinder, 1996; Egan, 1988; Egan & Gomez, 1985; Vicente, Hayes, & Williges, 1987).

The next subsection looks at the literature on spatial ability in more detail.

2.3 Spatial ability

Carroll (1993) identified eight categories of general intelligence. These categories included crystallized intelligence, fluid intelligence, general memory and learning, broad visual perception, broad auditory perception, broad retrieval ability, broad cognitive speed, and processing speed. Carroll categorized spatial ability as a factor in the broad visual perception category.

Benyon et al. (2000) reported that of all of the proposed components of intelligence, spatial ability has been the most frequently studied in connection with software use. Lohman (1996) suggested that spatial ability may be defined as "the ability to generate, retain, retrieve, and transform well-structured visual images" (p. 98). To provide a bigger picture of spatial ability this section addresses multiple aspects. Specifically, the remainder of this

subsection provides a description of how spatial knowledge is constructed, the differences related to egocentric and allocentric transformations, the influence of scale, and a presentation of common techniques used to assess these skills.

2.3.1 The construction of spatial knowledge

Thorndyke and Goldin (1981) performed a series of experiments that revealed that individuals with high visual skills were able to recall features in their neighborhood more accurately than individuals with low visual skills. In addition, subjects with high visual skills were also able to encode and manipulate mental representations of these features better. Mark, Freska, Hirtle, Lloyd and Tversky (1999) reviewed research on how people develop and work with geospatial knowledge in practical settings. Most researchers think that the construction of spatial knowledge can be conceptualized as consisting of three phases (Devlin & Berstein, 1995; Golledge, 1991; Siegel & White, 1975). In the first phase, a person is unfamiliar with an area. In this setting, the initial knowledgebase is declarative, existing in the form of landmarks, which are recognizable objects that are not yet attached to a mental map. During this early phase, the representation of landmark knowledge is usually in the form of a list, with little additional contextual information. As one becomes more accustomed to the layout of the area, procedural or route-based knowledge develops, in which routes are created in the form of connected steps between landmarks. During this phase, knowledge is represented as a linear sequence from an egocentric (self oriented) perspective (e.g., relying on terms such as right or left) and forms the beginnings of a network that is the basis for a mental map in which landmarks are embedded. Finally, as one accumulates geographic knowledge about the area, configuration or survey knowledge

develops. In this final phase, configuration knowledge is represented on a symbolic or figurative level using an external reference frame or coordinate system (e.g., cardinal directions), and leads to the creation of detailed cognitive maps.

2.3.2 Multiple dimensions

Lohman's (1996) definition of spatial ability proposed that this ability dealt with the capacity to perform multiple operations on visual images (i.e., generating, retaining, retrieving, and transforming). Because of its multiple dimensions, this definition would imply that spatial ability is not a unitary construct.

Specific to the interpretation of map content, the subset *orientation* construct of egocentric transformations is of interest. Orientation can take on one of two forms in either egocentric and allocentric (object-based) perspectives. Burgess (2006) suggested that these abilities exist in parallel. Egocentric perspective involves the ability to imagine a reoriented-self (perspective taking). Allocentric perspective involves the ability to imagine the rotation of objects from a fixed perspective (mental rotation). The following subsections provide details on differences related to egocentric and allocentric transformations, the influence of scale, example subset constructs of each perspective, and a presentation of common techniques used to assess these skills.

2.3.3 Egocentric perspective

Egocentric perspective involves sensory feedback that contributes to the sense of direction one possesses for the physical space surrounding them. This type of configuration is often involved with large scale interpretation (e.g., reading a map to find the way in a large city). This ability is comparable to the skill that would be required to make a judgment about

direction when using a map compared to the real-world. Hegarty, Richardson, Montello, Lovelace, and Subbiah (2002) and Hegarty, Montello, Richardson, Ishikawa, and Lovelace (2006) found a significant difference in types of spatial ability that exist for large-scale (for example, the information that a person receives about orientation when standing in an open space) and small-scale applications (for example, the spatial skills one would use to solve a rubik's cube). They found that large-scale spatial processing skills required sensory feedback.

Shepard and Metzler (1971) found that reaction time increases as the angular difference of orientation increases. Furthermore, people are faster and more accurate in naming and pointing at objects in front and behind them than those to the left or right when imagining perspectives (Bryant & Tversky, 1999; Hintzman, O'Dell, & Arndt, 1981). Kozhevnikov and Hegarty (2001) also found results on pointing accuracy where absolute angular error increased with the angular deviation of one's imagined heading from the orientation of the array.

A common technique used to measure spatial ability involves the use of psychometric tests. Hegarty et al. (2002) suggests that self-reported measures are a more promising approach to measuring large-scale environmental spatial ability than psychometric testing. Other techniques for measurement include sketch maps and perspective taking tests. Kozhevnikov, Motes, Rasch, and Blajenkova (2006) created a test known as the Perspective Taking Assessment to examine this ability. The test involves a computerized application where a display shows a person that is situated within an arrangement of landmarks. A person taking the test is told to imagine themselves in the position of the person in the display. In that arrangement the representative figure is facing in the direction of one of the

landmarks. After a few seconds a red flashing spot begins to flash on one of the other landmarks. The task of the person taking the test is to determine the pointing direction of that other landmark. After they have made that decision they must make a mental rotation and map the pointing direction onto a separate configuration of arrows in which the up arrow aligns with the forward view.

Two sub-factors of spatial ability that are commonly found significantly correlated with computer performance are orientation and visualization. Spatial visualization has been defined as the "ability to manipulate or transform the image of spatial patterns into other arrangements" (Ekstrom, French, Harman, & Dermen, 1976, p. 173). Spatial orientation has been defined as the "ability to perceive spatial patterns or to maintain orientation with respect to objects in space" (Ekstrom et al., 1976, p. 149). Furthermore orientation has been described in two forms. One form involves the mental rotation of objects. The other form involves perspective taking which is dependent on a person's frame of reference in relationship with the real world. These abilities have been shown to be distinct from one another yet most research on computer performance examines either one or the other (Pak, Rogers, & Fisk, 2006).

2.3.4 Allocentric perspective

In contrast to egocentric configurations, allocentric configurations involve "figural space" manipulations. Figural space has been defined as small in scale relative to the body and external to the individual (Hegarty et al., 2006). For example, solving the 2-d problems presented in the game Tetris.

Lohman (1996) and Carroll (1993) suggested that there are five constructs of spatial ability that include visualization, speeded rotation, closure speed, closure flexibility, and perceptual speed. Techniques that are commonly used for measuring spatial ability involve psychometric tests such as the paper-folding and number comparison tests which measure visualization and perceptual speed, respectively. Small-scale processing skills do not have large requirements for sensory feedback. Therefore when one would conduct a user study within a controlled lab the participants would be isolated to a small space and the confines of a computer display. Thus small-scale spatial ability would be the primary measurement of interest. Psychometric testing is often used as a paper-based measurement tool to assess spatial ability in user studies. A popular form of these tests is found in Ekstrom's Manual for Kit of Factor-Referenced Cognitive Tests (Ekstrom et al., 1976). This kit includes a diverse battery of cognitive tests that covers the constructs of spatial ability presented by Lohman (1996) and Carroll (1993).

The following subsections present detailed information on specific factors associated with allocentric perspective.

2.3.4.1 Visualization

Visualization is an aspect of spatial cognition that involves "ability in manipulating visual patterns, as indicated by the level of difficulty and complexity in visual stimulus material that can be handled successfully, without regard to the speed of task solution (Carroll, 1993, pp. 362-363)." McGee (1979) performed an analysis to understand the nature of visualization ability and concluded that it "involves a process of recognition, retention, and recall of a configuration in which there is movement among the internal parts of the

configuration or the recognition, retention and recall of an object manipulated in three-dimensional space" (p. 893). These descriptions of visualization support other research that proposed a relationship with the person's ability to make configurations. McGee's findings indicated that visualization is a reliable measurement for predicting success in vocational-technical training programs. Furthermore, based on a review of the literature, Sein, Olfman, Bostrom, and Davis (1993) suggested that there is a need to identify both effective training methods and user interfaces to accommodate users with low visualization ability. A common assessment used to measure visualization is the paper-folding test (Eliot & Smith, 1983). Sein et al. investigated visualization ability in relationship to usage of three applications: email, modeling software, and operating systems. They found that persons with high visualization skills learned fastest on all of the applications. They stress the importance of direct manipulation interfaces suggesting that this type of design helps the user to visualize system activities. In doing so, Sein et al. suggested that it alleviated the onus on the user to form a detailed internal (i.e., mental) representation. Thus users with low visualization skills can perform just as well as users with high visualization skills.

Pak et al. (2006) conducted research to compare the differences between two versions of a browser like interface. One version was presented with navigation in the form of a map structure and the other had navigation in a directional step-by-step layout. These systems were designed in such a way that the map-like interface was the most navigationally intensive. The results showed that spatial orientation ability was related to performance on tasks that required a high degree of navigation. There were no significant results that suggested a relationship with visualization. Their explanation for these findings was that the relationship with ability is dependent upon the nature of the task. In their case, because

navigation played an important role in completing the task, orientation ability was the best predictor. While this group made progress in testing both orientation and visualization they only tested on the mental rotation aspect of orientation which did not factor the person fitting a frame of reference.

2.3.4.2 Visual memory

Visual memory has been defined as "the ability to remember the configuration, location, and orientation of figural material" (Ekstrom et al., 1976, p. 109). Visual memory is not listed as a construct of spatial ability, however it is referenced as related to spatial visualization. Carroll (1974) has suggested that spatial visualization involves short-term visual memory.

A published test used for assessing visual ability is the Building Memory test from the Manual for Kit of Factor-Referenced Cognitive Tests (Ekstrom et al., 1976). Research has shown that visual memory carries over to how users interact with software interfaces (Ziefle & Bay, 2006).

2.3.4.3 Perceptual speed

Another construct of spatial ability is perceptual speed. This factor is a construct of spatial cognition that involves "speed in comparing figures or symbols, scanning to find figures or symbols, or carrying out other very simple tasks involving visual perception" (Ekstrom et al., 1976, p.123). A common assessment used to measure perceptual speed is a paper-based psychometric assessment known as the Number Comparison Test (Ekstrom et al., 1976).

Ackerman and Kanfer (1993) conducted a research study that tested air traffic controllers to see if there was a relationship between the different constructs of spatial ability and job performance. Their findings supported the broader spatial skills categorized by Lohman (1996) and Carroll (1993) in that they could be correlated with task performance, with the exception of perceptual speed. They suggested that it was not surprising that perceptual speed was not correlated because the task under evaluation involved procedures that were inconsistent and thus hard to achieve expertise at an automatic level of information processing.

While these studies did not show an effect of perceptual speed on computer performance, they were limited to applications that involved conventional labeling and layouts familiar to their users. When a person has knowledge of objects it is quite easy to distinguish between them. Given a set of unfamiliar objects however it takes time to build up an understanding of the differences between items. In that setting perceptual speed ability would influence user performance.

2.3.4.4 Reasoning

Logical reasoning was described by Carroll (1994) as involving retrieval of both meanings and algorithms from long-term memory that is followed by a performance of serial operations on the materials retrieved. Researchers have found working memory span to be related to reasoning (Ackerman, Beier, & Boyle, 2002) and have identified a set of characteristics and heuristics for this skill. Some of the popular reasoning tactics include reasoning by analogy, representativeness, availability, anchoring, framing, default, recognition, and taking-the-best approach.

Reasoning by analogy involves choosing a strategy where situations are very different, but there is an underlying deep similarity (i.e., an atom is like the solar system). The representativeness heuristic involves judging probability in terms of how well something represents, or matches a particular prototype. The availability heuristic involves judging the likelihood of events by their availability in memory. Anchoring effects occur because once a possible answer to a question is available, it is used as a reference point. For example, when presented with the question "How much longer is the Mississippi than 500 miles?" and "How much longer is the Mississippi than 100 miles?" people judge the Mississippi as longer when given 500 miles as an anchor. The framing factor suggests that the context in which you present an item can influence the way you judge that item. It can frame your attitude of something being positive as opposed to negative. For example, the following scenarios are the same: "out of 600, 200 people will live" or "out of 600, 400 people will die." Typically, the terminology used in each scenario frames the subject's mindset such that the first scenario is reasoned to be positive and the second to be negative. Another influential factor is the default heuristic. For example, organ donation is higher in France than in the US because it is the default choice. The recognition heuristic involves choosing an option because you recognize it and it seems reasonable that you recognize it (i.e., if you had to choose someone who is famous, you would go with a name you recognized because it is more likely that you remember the name because that person is famous). Another influential factor is the take-the-best heuristic. For this heuristic, if you recognize or are able to think of two different options, you will compare the two options by different dimensions until you find a difference and thus a more valid answer. This works when you have cues and knowledge about both choices and furthermore the knowledge is informative.

Allen (1994) investigated the relationship between perceptual speed and logical reasoning. He used an information retrieval system that presented information in two ways (rank-ordered and non-rank-ordered) to evaluate user performance. There was no effect of perceptual speed, however there was an interaction between logical reasoning and the user interface type. Allen concluded that users with low reasoning skills performed better using a rank-ordered system than a non-rank-ordered system.

While reasoning is not listed as one of the five constructs acknowledged by Lohman (1996) and Carroll (1993), it is widely accepted that the capacity to reason is related to spatial ability. Researchers have suggested that working memory capacity involves the ability to manipulate information and suppress distracting information (Engle, 2002). Other researchers have supported this idea and found working memory capacity to be related to reasoning (Ackerman et al., 2002). These abilities are important to spatial transformations because they help a person control attention and develop effective strategies to carry out rational complex mental rotations.

2.3.5 Spatial ability and computer performance

Vicente and Williges (1988) and Gomez et al. (1986) found that both people with high and low spatial ability benefit when an interface is tailored. When considering small screens on computers it is important to address the fact that the user will always have limited exposure to content because of space limitations. Thus the mental demands for recollecting information from background displays are high because there are many background screens. Woods (1984) conducted research on the effective distribution of attention when the user must process information across successive computer displays. He used the term "visual

momentum" to describe the measure of the user's ability to extract and integrate information across displays. He suggested that it is important to establish visual momentum to build a representative map of the underlying system or process that supports human spatial reasoning skills. Furthermore, if visual momentum is present it provides for a seamless interaction between the person and the computer.

The theoretical framework used to explain how spatial knowledge is constructed (i.e., landmark, route, and survey) has been found to be transferable to navigation in hypertext (McDonald & Stevenson, 1998; Vicente et al., 1987) and also for PDAs (Arning & Ziefle, 2006; Goodman, Gray, Khammampad, & Brewster, 2004; Ziefle & Bay, 2006). Ziefle and Bay investigated two types of navigation aids in a cell phone. The navigation aids included a category aid and a tree aid. The category aid involved a header that had subcategory links listed by descriptive text below it. The tree aid was identical to the category aid, however it also showed parents and parent-parents of the current category. The subcategories were indented to emphasize the hierarchical structure. The two aids were designed to convey landmark (category aid) and survey (tree aid) knowledge. The results of this study revealed that the tree aid had consistent and significant advantages for both young and old users. This advantage was larger for users with low spatial ability and older adults. They suggested that spatial information in a menu structure helped individuals with low spatial abilities because it conveyed survey knowledge that these individuals were slower to develop.

Büring, Gerken, and Reiterer (2006) conducted a study on small screens that tested the relationship of spatial ability and the effective use of a zoomable user interface. In the study, there were two forms of the interface: 1) a screen that included an overview window (showing a miniature of the entire information space) and 2) a screen that solely depended on

a detailed view (no overview window). The findings revealed that individuals who had lower spatial ability were helped by the overview window. In contrast, individuals with high spatial ability solved tasks faster without the overview. In fact, some participants with high spatial ability were actually slowed down when presented with the overview relative to low spatial ability users. These findings illustrate how individual differences in spatial ability influence the way in which people work with interfaces that have small screens.

2.4 Cognitive demands of spatial processing

The primary emphasis of this section revolves around cognitive demands on memory for spatial processing. Relevant literature and findings are presented in four sections including attention, theories for multiple memory systems, information processing, and working memory.

2.4.1 Attention

As a starting point, it is helpful to begin with attention because it involves the focal point for the aspects of memory that we are concerned with. A review of the cognitive science literature reveals that there are several theories about how attention should be defined. A frequently quoted definition of attention comes from classic work by William James. In his monumental Principles of Psychology (1890) he reported that "Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought" (p. 403-404).

Feature integration theory (FIT) proposes attention in the form of a spotlight (Wolfe, Cave, & Franzel, 1989) with limited capacity that processes information in a serial manner (Treisman & Gelade, 1980; Treisman, 1986). The spotlight metaphor suggests that attention can only be focused on one item or a group of items at a time and can only activate awareness of a certain amount of space. The car wash has been a popular analogy of this theory. For example, imagine a line of cars going through a car wash one at a time. It is only possible to wash (i.e., process) one car at a time. This model of attention suggested that the slope for processing information is determined by plotting reaction times as a function of the number of distracters. Distracters involve irrelevant information surrounding the target object. As the number of distracters increases, reaction time required prior to recognizing the target object also increases.

The FIT model proposed two ways by which individuals process visual information which involve pre-attentive and attentive processing. Pre-attentive processing allows for quick feature searches that aren't affected by the number of distracters. This form of visual processing is limited to a small set of feature differences such as color, size, motion, or orientation. For example, pre-attentive processing allows one to quickly spot a red item among several green items regardless of the amount of green items. Reaction time does not change in relationship to the number of distracters because there is only one feature difference that is required for object recognition. In contrast, attentive processing is required for conjunction searches.   Conjunction searches are performed when visual information must be bound for object recognition. For example, an object with disconnected joints (not conjoined) would be more challenging to find in a set of other objects with disconnected joins by comparison to a solid (conjoined) object.

Attention is a significant component of the aforementioned model because it represents the individual's focus of concentration. Kahneman's (1973) capacity model of attention suggests that there are two factors that influence a person's control over the allocation of attention. The first factor involves intention and experience. Specifically, the objects that the individual is more interested in and familiar with are more attractive. Applied to the FIT model, familiar objects have the potential to serve as distracters. The second factor involves evaluation of demands on capacity. This factor suggests that humans will evaluate the demands on capacity in their minds when there are multiple objects and usually give attention to the ones that require lower capacity.

2.4.2 Working memory

Baddeley and Hitch (1974) suggested that working memory consists of an attention director referred to a central executive in addition to three slave systems that include a phonological store, visuo-spatial sketchpad, and episodic buffer. The central executive controls and regulates the cognitive processes, directs attention, coordinates motor output, and controls the slave systems. The phonological store deals with sound and is the most understood component of working memory. The visuo-spatial sketchpad stores visual and spatial information (e.g., shapes, colors, location or speed of objects in space, etc.). The episodic buffer deals with conscious experiences. It leads to the registration of information to long-term memory.

Baddeley and Hitch (1974) are well known for their non-unitary model of working memory. Specific to this model, information is stored to short-term memory (STM) and transferred to long-term memory (LTM) as a result of encoding. This idea is typically

accepted because of the fact that when there is an overload on working memory the whole system does not crash. Craik and Lockhart (1972) proposed the idea that an implied hierarchy exists in processing where the strength of a memory trace depends on level of processing. They suggested that processing goes through levels from shallow to deep where deeper learning leads to longer lasting traces. The level of processing distinction explains data where two different forms of rehearsal occur (i.e., maintenance or elaborative). Maintenance rehearsal involves same-level rehearsal and does not increase LTM. It typically involves a rote repetition. Elaborative rehearsal involves encoding at a different level that moves toward deeper processing and thus increases LTM. Elaborative rehearsal typically involves a more meaningful strategy of encoding than maintenance rehearsal and results in a better transfer to LTM. For example, mnemonics to remember sets of words or objects. An applied incidental memory study that supports these ideas came from Craik and Tulving (1972). They investigated three conditions: 1) physical (e.g., the subject was asked "Does the word DEER contain and E?"), 2) acoustic (e.g., the subject was asked "Does the word DEER rhyme with fear?"), and 3) semantic (e.g., the subject was asked "Is the DEER a living thing?"). The results showed that encoding was best for the semantic, then the acoustic, then the physical. The level of processing view explains these results where deeper processing was occurring for the semantic condition and to a lesser degree the acoustic condition.

2.4.3 Multiple memory systems

There is a widespread argument in the field of psychology for multiple memory systems. Specifically, researchers argue that there are different systems to support various types of memory (e.g., explicit, implicit, declarative and procedural memory). Explicit

memory is the conscious recollection of past experiences (i.e., observing or hearing an item or event). Explicit memory is often tested through free recall, cued recall, and recognition tests. Free recall involves having a subject recall a comprehensive list of events or items at the very end of a test session. Cued recall involves having a subject recall items or events at certain intervals throughout a testing session. Recognition tests involve having a subject identify target items from a list of many possible choices (e.g., multiple choice tests).

Declarative memory is knowledge that involves factual information that is stored up in conscious memory. For example, "George Washington was a former president of the United States" is a fact that would be stored in declarative memory. Procedural memory is knowledge expressed in our behaviors that we are not conscious of. An example that illustrates this well is the rules that control our speech. To clearly express an idea there are grammatical rules that are used to structure sentences. While we don't step through each of those rules every time we speak, they are running in the background subconsciously. Declarative memory influences a person's familiarity with conventional labels. Conventional labels might include commonly accepted graphics and/or text used to display menu items. In this case, examples of declarative memory in storage would include items such as images remembered for the "save" and "paste" icons. In contrast, users would have a lack of declarative memory for non-conventional labels. Non-conventional labels are associated with functions that are not commonly used and commonly specific to a special type of rare job. Therefore, these labels would involve a bigger learning curve because new declarative memories would have to be developed.

Similar to spatial ability, it is commonly accepted that working memory is not unitary (Baddeley and Hitch (1974). That is, as opposed to there being one type of memory there are

multiple memory system types. Tulving (1985) used experiments to test cognitive processes occurring in multiple memory systems. He found that a particular lesion or type of stimulation that affects the performance on one task but not on another is suggestive of the existence of multiple memory systems. Tulving's research investigated performance on two tasks, word recognition and fragment completion. Word recognition involves presenting a subject with a list of words to remember and then presenting them with a different list after memorization in which they are tested to see how many of the words they can successfully recognize. Word fragment completion involves presenting a subject with words that are not fully completed (e.g., "h_gh_ay"). The subject is tested to see how many partial words s/he can complete correctly. A comparison of these tasks would suggest that they should involve different skills and memory retrieval. Tulving found just that where the level of performance between the two tasks was not correlated (or stochastically independent). Evidence in this case he argues is compelling because it was within-subject and thus he suggested it is strong support for cognitive processes occurring in multiple memory systems. Tulving's (1985) research can be applied to examine the differences between spatial visualization and orientation (i.e., perspective taking). Much of the current literature supports the notion that these two skills are dissociable from one another (Pak et al, 2006). Extending Tulving's work into this context would suggest that if there is a lack of correlation within a model that includes both these skills and performance it is likely that different memory systems are in play.

2.4.4 Information processing

In the context of skill acquisition, Ackerman (1987, 1988, 1990, 1992) suggests that cognitive effort can be evaluated throughout a series of three learning stages. The model incorporates concepts from work by Fitts and Posner (1967) and the stages include a cognitive stage, associative phase, and an autonomous phase. The cognitive stage is the first phase and thus involves large demands in terms of the cognitive effort involved in understanding instructions and formulating strategies. Second, the associative phase involves proceduralizing task strategies that are established from the cognitive stage to enhance performance and reduce errors. The third stage is the autonomous phase where the task becomes automatic. In the final autonomous phase, the refinement process is completed and optimal strategies are all in place. This skill acquisition model predicts that as a user enters into a more stable state, they can work on the task with a lower level of consciousness. Performance in this final phase reaches a plateau and resembles results concurrent with the power law of practice effects.

Ackerman's skill acquisition model (1987, 1988, 1990, 1992) is relevant to button label and layout when usage is considered in a repetitive task. When a user first begins to use an application a period of training occurs. As the user first steps through the cognitive stage they are still acquainting themselves with the label meanings and their relative positions on the screen. In this phase, the user's retention for each label is quite limited because they have not internalized a strategy for recognizing the meaning of the label and its relative position. As the user progresses to the next associative stage of acquisition, they have identified ways to remember what buttons mean and also have a general mental model

for the layout of the screen. Thus their response time improves, however they are still refining strategies to optimize performance.

When allocating attention to these small levels of thought researchers have proposed different levels of processing. Atkinson and Shiffrin (1971) presented a model for information processing that involves three memory stores including sensory, short-term, and long-term memory stores. Each of these stores has unique characteristics. The sensory store involves encoding of sensory information, storage that is typically less than one second in duration, a large capacity, retrieval where attention is needed, and forgetting that results from decay or loss. The short-term store involves encoding of acoustic information, storage that is typically 15-30 seconds in duration, limited capacity (7 ± 2 chunks), retrieval that is subjectively instantaneous, and forgetting that results from distracters (interference) and information loss. The long-term store involves encoding of semantic (meaningful) information, storage that is typically greater than 30 seconds in duration, unlimited capacity, retrieval where there is an active search, and forgetting that results from interference and retrieval failure.

Shiffrin and Schneider (1977) proposed a model of automatic and controlled processing for how people control their attention to process information and suggested these modes are influenced by the previously mentioned stores. Automatic processing is fast, parallel, not limited by short term memory, and not under direct control. Because of the lack of direct control it is difficult to suppress or alter. Furthermore it requires consistent mapping to occur. Consistent mapping involves stimuli roles that never change across trials. That is, an item classified as a target would always be a target and an item classified as a distracter would always be a distracter. The subject is able to make the same overt/covert response

across trials and thus the task has the potential for becoming automatically processed. In contrast to automatic processing, controlled processing is slow, generally serial, effortful, capacity limited, and used when dealing with novel or inconsistent information. Performance improves because the control process is changed or because automatic processing develops. While automatic processing can only be established under consistent mapping conditions, controlled processing can handle varied mapping. Varied mapping involves stimuli that change roles across trials. That is, an item might be specified as a target or distracter. This type of condition causes interference and thus automatic processing will not occur.

2.4.5 Working memory span

Engle (2002) proposed that differences in working memory span (or capacity) can be explained by assessing a person's ability to control attention to maintain information in an active and quickly retrievable state. Engle suggested that working memory span is an important individual difference. Researchers have recently suggested that working memory is a major predictor of general intelligence (commonly referred to as *g)* (e.g., Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen 2004; Ackerman et al., 2002). Vernon (1950) suggested that *g* accounts for 40% of the total variance in human abilities.

Working memory span is often measured through use of reading, operation, digit, and counting span tasks. Each of these tasks involves presenting subjects with a series of items that are distributed with distracters. For example, the reading span task involves having the subject read a series of two to seven sentences that are followed by an unrelated word. At the end of the session the subjects are asked to recall all of the unrelated words. In this case, the two to seven sentences serve as the distracters. Essentially the idea is to tap the subject's

resources for doing something in addition to encoding recall items. Researchers have found working memory span to be indicative of how well people perform on the Stroop task (Engle, 2002). Kane and Engle (2003) found that persons with high working memory span did better than low span persons only when at least 75% of the items were incongruent.

2.5 Guided applications

A common principle suggested in the field of usability is to minimize user memory load (Nielsen, 1994). The premise is that by doing so, users are relieved of the needs on their working memory so that they can allocate the majority of their cognitive resources on the task at hand. A common response to this guideline is to assume that providing assistance through visual cues is the answer, however just the opposite has been discovered to be true. Researchers have found that an unguided interface resulted in more efficient performance than a guided interface (van Nimwegen & Oostendorp, 2007). The explanation for these findings is that the unguided interface involves more active thinking and contemplation. This idea coincides with Craik and Lockhart's (1972) level of processing theory that deeper processing will result in deeper encoding.

The differences between guided and unguided systems can be understood through recognizing two forms of how information processing can occur. First, guided systems involve an externalization of information. Externalized information is often in the form of visual cues that lead the user to choose an appropriate response or action. For example, wizards, help-options, and menu options that are grayed out are common design features that are in place to help users. These forms of externalized information provide a means to relieve working memory because the relevant information reduces demands for recall.

Externalization therefore involves the display of guidance information that could be useful to working through the current task. The challenge in designing for effective externalization is that it has the potential of two negative effects. First, if too much information is provided it creates demands on working memory to interpret the excessive information. Second, it decreases the demands on exploratory learning. Therefore because the user does not go through a personal process of learning through exploration they lack a deep knowledge of the system in the end. Zhang (1997) conducted a study on externalization and found that external representations could constrain, structure, and change cognitive behavior. Furthermore, he found that redistributing information from internal memory to an external display improved performance. Zhang concluded that externalization can be beneficial if the benefit of using external representations can offset the cost associated with the externalization process.

In unguided systems, information or clues are not readily available for assistance. Therefore information must be internalized or inferred and stored in memory prior to usage. The benefit of internalization is that it requires the user to go through an exploratory learning period. Therefore recall of unguided systems is typically better than the recall that occurs with guided systems. The limitation on these systems exists for persons with lower abilities. Researchers have commonly found that when designing computerized systems for students learning by discovery is generally observed to provide greater gains for students with high ability and greater losses for students with low abilities (Berger, Lu, Belzer, & Voss, 1994).

In summary, unguided systems can be beneficial to persons with high abilities where they will lead to a strong memory of the interface. Guided systems often are more useful than unguided systems for persons with low ability. Guided systems could also be just as

useful to persons with high abilities so long as there is an optimal amount of information. Dependent on the type of individual difference, specific characteristics would influence the way in which users internalize information. Zhang and Norman (1994) argue that when considering the representation of a distributed cognitive task it is never solely internal or external. Instead they propose a system of distributed representations with both internal and external indispensable parts. Therefore both individual differences and system design whether guided or unguided have the potential to influence computer performance.

2.6 Summary

The literature presented in this chapter highlighted relevant findings to our interests in investigating the role of spatial ability in interface design. Specifically, individual differences, spatial ability and reported significant relationships with computer performance, cognition, and guided applications. The following chapters present two research studies that synthesize many of these ideas into applied settings.

CHAPTER 3. THE ROLE OF SPATIALLY RELATED COGNITIVE ABILITIES IN USING A SIMPLE PEN-BASED INTERFACE

Abstract. Researchers in human computer interaction (HCI) have recognized the importance of individual differences in cognitive abilities for designing effective software interfaces. Spatial ability has been shown to be one of the most important individual differences affecting user performance at the global level, but little is known about its impact on performance in relation to button layouts or the salience of the button label. We investigated the relationship between spatial ability and user performance in the context of simple interfaces that involve selecting a button to match a specified task. Subjects were given three layouts that included a circle of twelve buttons associated with twelve different colors, two horizontal rows with the twelve buttons linked to colors, and two horizontal rows associated with functions on a map software interface. Subjects were randomly assigned to one of four button label treatments: informative graphical labels, informative text labels, uninformative arbitrary graphic labels, and uninformative no button labels. We hypothesized that subjects with high spatial visualization ability would perform better in selecting the correct software button for a specified task than subjects with low spatial visualization ability, and that differences between users with low and high spatial ability would increase with the more complex map interface and with button labels that were uninformative. We found that on average, subjects with high spatial visualization ability required significantly less time to select the correct button at the beginning of each exercise and were better able to correctly recall button location after a series of exercises for the interface. We also found that on average, subjects of high spatial visualization took less time to correctly identify

buttons representing map operations, which involved software buttons that are not commonly used in day to day tasks.

3.1 Introduction

Researchers in human computer interaction (HCI) have recognized the importance of individual differences in cognitive abilities for designing effective software interfaces. Dillon and Watson (1996) argue that this approach is needed because traditional methods, such as basing interface design on generic models of users or giving all users standardized training, are not sufficient for addressing the variation that exists in users. Egan (1988) found that user performance in accessing information via software interfaces is affected by an individual's spatial ability, memory, reasoning ability, verbal aptitude, and possibly personality. Of these abilities, spatial ability has been found to be one of the strongest predictors of human computer performance (Dahlbäck, Höök, & Sjölinder, 1996; Egan, 1988; Vicente, Hayes, & Williges, 1987, Egan & Gomez, 1985). Lohman (1996) defined spatial ability as "the ability to generate, retain, retrieve, and transform well-structured visual images" (p. 98). The relationship between user performance and spatial ability is believed to arise from the fact that interacting with an interface is by nature spatial. For example, in accomplishing a software task, one may search the interface to find appropriate buttons or navigate through tiers of hierarchical menus (e.g., Vicente et al., 1987). The theoretical framework that explains how spatial knowledge is constructed has been found to be transferable to navigation in hypertext (Chen & Rada, 1996; Dahlbäck et al., 1996; McGrath, 1992) and also for PDAs (Arning & Ziefle, 2006; Goodman, Gray, Khammampad, & Brewster, 2004; Ziefle et al., 2006). Woods (1984) conducted research on the effective

distribution of attention when the user must process information across successive computer displays. He used the term "visual momentum" to describe the measure of the user's ability to extract and integrate information across displays. He found that it was important to establish visual momentum to build a representative map of the underlying system or process that supported human spatial reasoning skills. He suggested that if visual momentum was present it provided for a seamless interaction between the person and the computer. Due to the limitations of presenting information on small screens, handhelds inherently involve many successive screens. Visual momentum therefore would play a significant role in a person's ability to effectively develop spatial knowledge of handheld applications.

These results suggest that a design approach that accommodates individual differences, especially spatial ability, is critical to creating handheld computer interfaces for displaying maps to support tasks in outdoor settings. We are motivated by the need to design software to facilitate field tasks for large-scale statistical surveys and censuses. Field tasks include canvassing streets to list addresses for sample selection, or navigating through neighborhoods to find households selected for interviews. In this setting, a large and diverse field staff uses handheld computer software that presents maps for listing, finding or updating addresses. Because of the inherently geospatial nature of navigation and address verification tasks, we expect spatial ability to be an important factor in the effectiveness of the interface, particularly given the widely varying abilities to work with geospatial information in this user population. In addition, map-based software functions and their icons are unfamiliar to many people, even if they commonly use email, browser or word processing software. Thus, learning icon labels or button locations might be more challenging for novice users, especially those with low spatial ability.

To create a research setting, we looked at Ehret's (2000, 2002) study that investigated the ability to learn different software button locations in relation to whether or not software button labels were related to or were unrelated to the button's software function. He found that users were better able to remember the location of buttons with labels that were uninformative compared to those that had an informative label, in part because the cognitive effort required to recall these buttons was much greater than for buttons with informative labels. While the emphasis on button relevance and learning the location of buttons is appropriate for our motivating application, Ehret's work deviated somewhat from our desired setting. First, the button layout was circular, which is not a standard layout in current software design. Second, the interface did not include the level of complexity more typical of map software. Finally, no information was gathered on individual differences.

Our goal was to evaluate the association between spatial ability and user performance in a setting that included the more complex content and button functions associated with map software. A study was conducted that investigated user learning of button locations in relation to the relevance of the button label to its function (informative graphic, informative text, arbitrary graphic, no label) and the arrangement of the buttons into an interface layout (Ehret's circular button layout, grid layout, grid layout with map function buttons and display). We hypothesized that subjects with high spatial ability would have better performance than subjects with low spatial ability. We also hypothesized that performance differences for subjects of low and high spatial ability would be smaller with informative software button labels than with uninformative labels and smaller for simple interfaces relative to an interface with map button functions and content.

The following sections present a background on individual differences, especially differences in spatial ability, and the influence of working memory. We then outline methods used in our study, the results from the study, and discuss its implications.

3.2 Background

3.2.1 Individual differences

Benyon et al. (2001) suggested that individual differences are an important distinction to consider across computer users because of the impact they have on the ability to derive strategies to interpret the behavior of technological devices. HCI is a field that typically concerns the expression of abstract symbols (e.g., button labels) that represent something else (e.g., software functions). Thus because the end user is presented with information in the form of symbols and most often are restricted from seeing the underlying design of the system, they are left to their own devices (i.e., cognitive abilities) to interpret machine behavior. Cognitive abilities influence a person's capacity to develop the essential strategies necessary for interpretation. For this reason, Benyon et al. stressed the importance of considering individual differences in the area of HCI.

Benyon et al. (2001) presented three categories for individual differences that include physiological, socio-cultural, and psychological characteristics. Physiological differences involve biological classifications such as gender, size, vision, hearing, and mobility. Socio-cultural differences involve characteristics related to language, cultural, and environmental aspects of an individual's surroundings. The limitation on investigating physiological and socio-cultural characteristics for software design is that they involve measures that are typically too broad to be useful. For example, examining gender provides a classification of

reasoning ability, verbal aptitude, and possibly personality. While Dillon and Watson (1996) reported that personality differences have shown little support in predicting performance on user interfaces, the other abilities suggested by Egan offer potential. Of these abilities, spatial ability has been found to be one of the strongest predictors of human computer performance (Dahlbäck, Höök, & Sjölinder, 1996; Egan, 1988; Egan & Gomez, 1985; Vicente, Hayes, & Williges, 1987). We will discuss spatial ability and related abilities in the following sections.

3.2.2 Spatial ability

Lohman (1996) suggested that spatial ability may be defined as "the ability to generate, retain, retrieve, and transform well-structured visual images" (p. 98). This definition implies that spatial ability is not a unitary construct. Lohman (1996) and Carroll (1993) suggested that there are five constructs of spatial ability that include visualization, speeded rotation, closure speed, closure flexibility, and perceptual speed. Two of these constructs, visualization and perceptual speed, have often been cited in the HCI literature as being related to computer performance. Speeded rotation requires a mental rotation of an object. Both closure speed and flexibility are concerned with the speed of apprehending and identifying visual patterns, often in the presence of many distracting stimuli. While speeded rotation, closure speed, and closure flexibility offer insight into how well a person will effectively use an application we did not recognize any direct parallels. The task of interest did not involve mental rotation nor did it include excessive clutter.

3.2.2.1 Visualization

Visualization is a construct of spatial cognition that involves "ability in manipulating visual patterns, as indicated by level of difficulty and complexity in visual stimulus material that can be handled successfully, without regard to the speed of task solution" (Carroll, 1993, pp. 362-363). Others have emphasized that a unique aspect of visualization has to do with the way it involves restructuring components within the visual stimulus in order to make manipulations. This restructuring would require a high demands on the control of thought in working memory due to the serial operations involved to perform each task. Carroll (1974) suggested that visualization involves a mental rotation of a spatial configuration in visual short-term memory. McGee (1979) performed an analysis to understand the nature of visualization ability and concluded that it "involves a process of recognition, retention, and recall of a configuration in which there is movement among the internal parts of the configuration or the recognition, retention and recall of an object manipulated in three-dimensional space" (p. 893). These definitions support other research that proposes a relationship with the person's ability to make configurations.

Sein, Olfman, Bostrom, and Davis (1993) investigated visualization ability in relationship to usage of three applications (email, modeling software, and operating systems). They found that subjects with high visualization skills learned fastest on all of the applications. With appropriate training and interface design, the gap between high and low visualization users was reduced and even reversed. Visualization is often significantly correlated with computer performance (e.g., Vicente et al., 1987). Miyake, Friedman, Rettinger, Shah, & Hegarty (2001) suggested that visualization requires the highest level of executive involvement (i.e., control of working memory). When compared to other aspects

of spatial ability (e.g., spatial orientation), visualization involves not only mental rotation but also re-mapping of parts. Thus, it is not surprising that visualization is frequently found to be correlated with computer performance and training because the thought processes associated with these tasks are often comparable with the cognitive complexities involved with visualization. While spatial visualization is commonly correlated with computer performance it is important to note that these findings are typically linked with complex tasks. Pak, Rogers and Fisk (2006) suggested that spatial abilities may be more highly associated with user performance when task difficulty is high but less so when task difficulty is low.

Psychometric testing is often used to assess spatial ability in user studies. A popular form of these tests is found in Ekstrom's Manual for Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Dermen, 1976). This kit includes a diverse battery of cognitive tests that covers the constructs of spatial ability presented by Lohman (1996) and Carroll (1993). One of the tests for evaluating visualization abilities is the Paper Folding Test. Shepard and Feng (1972) described the mental processes involved in paper-folding activities. They found that the number of serial operations (represented by individual folds), the number of relevant surface components carried along for each fold, and the number of irrelevant surface components influenced reaction time. This study showed that the extraneous information contributed to the burden on short-term memory and thus processing time.

3.2.2.2 Visual memory

Visual memory has been defined as "the ability to remember the configuration, location, and orientation of figural material" (Ekstrom et al., 1976, p. 109). A published test used for assessing visual ability is the Building Memory test from the Manual for Kit of Factor-Referenced Cognitive Tests. The test involves memorizing building locations on a map. This test involves strategies that would be used to learn button locations, and thus we recognized it as an ability that could potentially be correlated with task performance.

3.2.2.3 Perceptual speed

Another construct of spatial ability is perceptual speed. This factor is a construct of spatial cognition that involves "speed in comparing figures or symbols, scanning to find figures or symbols, or carrying out other very simple tasks involving visual perception" (Ekstrom et al., 1976, p.123). It influences a person's ability to distinguish between objects. A common assessment used to measure perceptual speed is the Number Comparison Test (Ekstrom et al., 1976).

Ackerman and Kanfer (1993) conducted a research study that tested air traffic controllers to see if there was a relationship between the different constructs of spatial ability and job performance. Their findings supported the broader spatial skills categorized by Lohman (1996) and Carroll (1993) where they could be correlated with task performance, with the exception of perceptual speed. They suggested that perceptual speed was not correlated because the task under evaluation involved procedures that were inconsistent and thus it was hard to achieve expertise in at the automatic level. Thus, there was not a valid correlation in performance because their findings matched their theory that measures of

perceptual speed are related to the performance of tasks that are automatized. In another study, Allen (1994) investigated the relationship between perceptual speed and logical reasoning with performance using an information retrieval system that presented information in two ways (rank-ordered and non-rank-ordered). There was no effect of perceptual speed, however there was an interaction between logical reasoning and interface type.

While these studies did not show an effect of perceptual speed on computer performance, they were also limited to applications that involved conventional labeling and layouts familiar to their users. Perceptual speed depends upon a person's sensitivity to distinguishing contrasts that exist amongst objects. When a person has an existing knowledge of objects it is quite easy to distinguish between them. When presented with unfamiliar objects, however, it takes time to learn to distinguish among items. Therefore, in learning a new application (i.e., especially those in which the user would not have any preexisting knowledge), perceptual speed ability may have an influence on performance with that application.

3.2.2.4 Spatial construction

The memory demands of spatial visualization and perceptual speed can also play a role on how people traverse space. Mark, Freska, Hirtle, Lloyd and Tversky (1999) reviewed research on how people develop and work with geospatial knowledge in practical settings. Most researchers think that the construction of spatial knowledge can be conceptualized as consisting of three phases (Devlin & Berstein, 1995; Golledge, 1991; Siegel & White, 1975). In the first phase, a person is unfamiliar with an area. In this setting, the initial knowledgebase is declarative, existing in the form of landmarks, which are recognizable

objects that are not yet attached to a mental map. During this early phase, the representation

of landmark knowledge is usually in the form of a list, with little additional contextual

information. As one becomes more accustomed to the layout of the area, procedural or

route-based knowledge develops, in which routes are created in the form of connected steps

between landmarks. During this phase, knowledge is represented as a linear sequence from

an egocentric (self oriented) perspective (e.g., relying on terms such as right or left) and

forms the beginnings of a network that is the basis for a mental map in which landmarks are

embedded. Finally, as one accumulates geographic knowledge about the area, configuration

or survey knowledge develops. In this final phase, configuration knowledge is represented

on a symbolic or figurative level using an external reference frame or coordinate system

(e.g., cardinal directions), and leads to the creation of detailed cognitive maps.

This construction of spatial knowledge concept has been a popular perspective on the

construction of spatial knowledge in the physical environment. Other researchers have found

this construction of spatial knowledge applies to computer usage. One example comes from

Ziefle and Bay (2006) who investigated two types of navigation aids in a cell phone. The

navigation aids included a category aid and a tree aid. The category aid involved a header

that had subcategory links listed by descriptive text below it. The tree aid was identical to

the category aid, however it also showed parents and parent-parents of the current category.

The subcategories were indented to emphasize the hierarchical structure. The two aids were

designed to convey landmark (category aid) and survey (tree aid) knowledge. The results of

this study revealed that the tree aid had consistent and significant advantages for both young

and old users. This advantage was larger for users with low spatial ability and older adults.

They suggested that spatial information in a menu structure helped individuals with low

spatial abilities because it conveyed survey knowledge that these individuals were slower to develop.

3.2.3 The influence of working memory on location learning

Ehret's (2000, 2002) primary hypothesis was that users would learn and retain the location of buttons with uninformative labels (e.g., arbitrary or no label) more quickly than button styles that required minimal effort to evaluate (e.g., informative graphic or text). The results from his study supported this idea. Differences in subject's ability to learn button locations may have been related to differences in how the memory of a button location was stored (or encoded) in working memory across treatments.

Engle (2002) suggested that working memory capacity is an important individual difference.   Researchers have recently suggested that working memory is a major predictor of general intelligence (commonly referred to as *g)* (e.g., Ackerman, Beier, & Boyle 2002; Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004). Vernon (1950) suggested that *g* accounts for 40% of the total variance in human abilities. Baddeley and Hitch (1974) suggested a non-unitary model of working memory. Furthermore, Baddeley (1986) defined working memory as "the temporary storage of information that is being processed in any range of cognitive tasks" (p. 34). Specific to this non-unitary model, information is stored to short-term memory (STM) and transferred to long-term memory (LTM) as a result of encoding.

Atkinsin and Shiffrin (1971) proposed that the main mechanism of forgetting information from LTM is the inadequate selection of probes (search cues) to be used for retrieval. Search probes are often created as the result of rehearsal. Craik and Lockhart

(1972) suggested that the type of rehearsal has an effect on the transfer rate of memory from STM to LTM. Two specific types of rehearsal commonly suggested are maintenance and elaborative rehearsal. Maintenance rehearsal typically involves a sub-vocal repetition. In contrast, elaborative rehearsal typically involves a more meaningful strategy of encoding that result in a better transfer to LTM than maintenance rehearsal. For example, use of mnemonics as a strategy to remember sets of words or objects, such as "Every good boy does fine" to remember line notes in sheet music and "FACE" to remember the space notes. Craik and Lockhart proposed the idea that an implied hierarchy exists in processing where the strength of a memory trace depends on level of processing. They suggested that processing goes through levels from shallow to deep where deeper learning leads to longer lasting traces. Thus where elaborative rehearsal involves a deeper level of encoding, recall is often found to be better than that found for maintenance rehearsal.

Ehret (2000, 2002) found that subjects with uninformative labels (i.e., arbitrary and no label) had better recall for button location than subjects who were given meaningful labels (i.e., informative graphic and text). Craik and Lockhart's (1972) levels of processing view offer an explanation for these findings. They proposed that stimuli are processed to different levels ranging from shallow to deep. This view implies a hierarchy where deeper learning leads to longer lasting memory traces that in turn increase the probability of its presence in LTM.

Another question that arose for the arbitrary condition was whether or not it presented the potential for incongruent associations comparable to the Stroop task. The Stroop task investigates responses to items with incongruent features (e.g., presenting the word red in a green font color). Researchers have found working memory span to be indicative of how

well people perform on the Stroop task. Engle (2002) proposed that differences in working memory span can be explained by assessing a person's ability to control attention to maintain information in an active and quickly retrievable state. In Ehret's (2000, 2002) study, one might have expected a difference between subjects with high and low working memory span, where incongruent terms existed in the arbitrary condition. For example, the graphic of a plane was used as one of the labels in Ehret's study (for a complete list see the Arbitrary Circle row in Figure 2). If the plane was not associated with the color gray it may have caused a mismatch. Although most of these graphics were neutral to having a preconceived color already associated with it, several of the labels might have had incongruence with their assigned color. Kane and Engle (2003) found that persons with high working memory span did better than low span persons only when at least 75% of the items were incongruent. Because only a few labels might have caused incongruence within Ehret's experiment (i.e., incongruent items < 75%) we did not consider incongruence to be a confounding factor.

3.3 Method

3.3.1 Overview

We conducted a study to evaluate whether spatial ability affected performance when using different types of button labels and interface designs. We recruited 78 subjects to perform 3 sets of exercises, each with a different kind of interface. Each participant was randomly assigned a button label treatment that varied from informative to uninformative labels. Software was developed to guide the subjects through a series of simple tasks on each of the interface layouts. The application recorded time to perform the task, number of attempts to complete the task, and the error in a users' recall of button locations. Because the

experiment had a repeated measures design we investigated the potential relationship between cognitive test scores and performance with a mixed model analysis of covariance.

3.3.2 Participants

The 78 subjects were recruited through fliers posted on the Iowa State campus, local grocery stores, retirement communities, in addition to newspaper ads and word of mouth. This method was used because we wanted to obtain data from a broad range of subjects. Prior to advertising, ten categories were specified which were used for recruiting subjects. These categories were: 1) female undergraduate student and 18-24 years of age, 2) male undergraduate student and 18-24 years of age, 3) female graduate student or non-student and 18-35 years of age, 4) male graduate student or non-student and 18-35 years of age, 5) female 36-50 years of age, 6) male 36-50 years of age, 7) female 51-65 years of age, 8) male 51-65 years of age, 9) female 66+ years of age, and 10) male 66+ years of age. Upon first contact, subjects were asked which gender and age category applied to them. Our goal was to include eight subjects within each category. Although 78 subjects participated in the experiment, only the data collected from 63 individuals were used because some elderly subjects were not able to complete the full experiment. The demographic distribution of the analysis data set included eight female undergraduate students 18-24 years of age, eight male undergraduate students that were 18-24 years of age, eight females classified as a graduate student or non-student 18-35 years of age, eight males classified as a graduate student or non-student 18-35 years of age, seven females 36-50 years of age, six males 36-50 years of age, seven females 51-65 years of age, seven males 51-65 years of age, two females 66+ years of age, and two males 66+ years of age.

49

Additional demographic and computer experience information was collected at the time of the study. Twenty-five percent were undergraduate students, 13% were graduate students, and 62% were not students. Forty-seven percent were employed full-time and 40% part-time, 5% were neither employed nor retired, 5% were retired, and 3% specified some type of other work status. Seventy-eight percent of the subjects reported having a high level of experience using a personal PC or laptop computer. Ninety percent reported a high level of experience using email, 81% working with word processing applications, 79% with web browsing, and 48% playing games. Only 30% reported a high level of experience using map software and 6% using global positioning system (GPS). Twenty-nine percent reported a high level of experience using drawing/graphics applications. Ninety-five percent of all subjects reported that they had little to no experience using a tablet PC. Seventy percent reported little to no experience using handheld computers.

3.3.3 Software

The interface consisted of buttons that corresponded to 12 colors or 12 map related functions (Figure 1). There was also a rectangular target box on the screen in which a randomly selected target appeared to signify the type of button to be selected. In addition, the third map exercise included a map display. Each trial started out with a screen that consisted only of the rectangle target box. After tapping on this box, a representation of the 12 color or map related functions would appear. The main task then involved selecting the button from a consistently mapped layout that matched the target on the screen. After selecting a button, if correct, all 12 buttons would disappear and the interface would be reset with a new target color in the rectangle box. If the selected button was not correct, a dialog

box would appear that told the user their selection was not correct and it instructed the user to
try again.

Four button label treatments were used as shown in Figures 2 and 3. The four
alternative button label treatments were: (a) a informative graphic that depicted the target
function associated with the button (for example, the button with the light green dot was the
correct match when the rectangle in the center of the screen was light green); (b) informative
text that described the target function associated with the button; (c) an arbitrary graphic that
was irrelevant to the target function associated with the button; and (d) no label on the
button. Each participant was assigned one of the four treatments which they in turn used
with all three of the layouts.
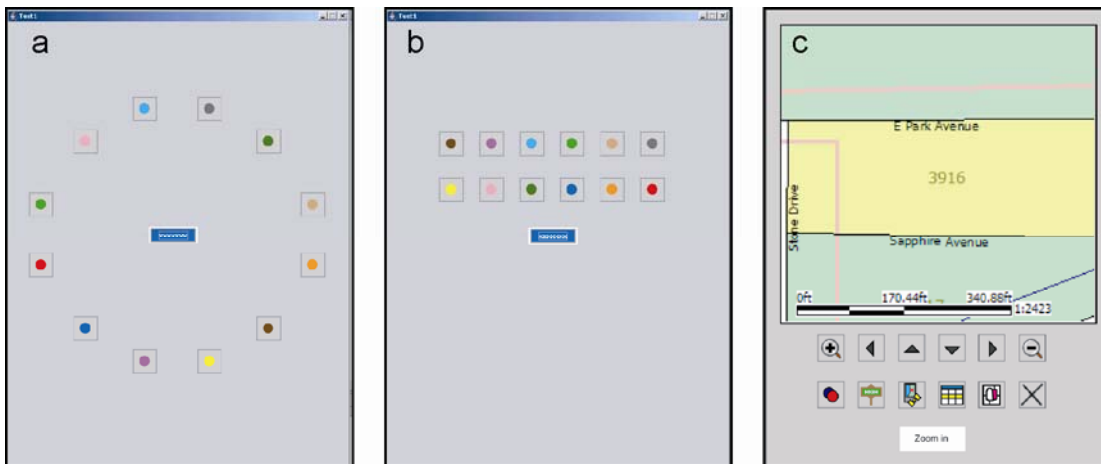
Figure 1: The three layouts used in the study shown with the informative button labels.

Treatment	Button Function											
Code	1	2	3	4	5	6	7	8	9	10	11	12
Informative graphic	▶	◀	▲	▼	zoom in	zoom out		HIGH				×
Informative text	Pan right	Pan left	Pan up	Pan down	Zoom in	Zoom out	Assign. area	Street names	You are here	Street list	Inter- sections	Close map
Arbitrary		⊘	◉	car			bulb	♠			◆	
No label												

Figure 2: Button labels for the circle and grid exercise.

Treatment	Button Function											
Code	1	2	3	4	5	6	7	8	9	10	11	12
Informative graphic	▶	◀	▲	▼	zoom in	zoom out		HIGH				×
Informative text	Pan right	Pan left	Pan up	Pan down	Zoom in	Zoom out	Assign. area	Street names	You are here	Street list	Inter- sections	Close map
Arbitrary		⊘	◉	car			bulb	♠			◆	
No label												

Figure 3: Button labels for the map exercise.

Exercise one used a circular layout of 12 buttons where each button was associated

with a unique color (Figure 1a). The target rectangular box was filled with one of the 12

possible colors. The task was to tap on the button that corresponded to the target color. The

color associations for the arbitrary and no label buttons were learned through the use of

tooltips. In this case, hovering over a button generated a tooltip of the associated color

(Figure 4). Exercise two used a grid layout that was the most of typical of software button

configurations (Figure 1b). The main task was the same as the one performed for exercise

one.
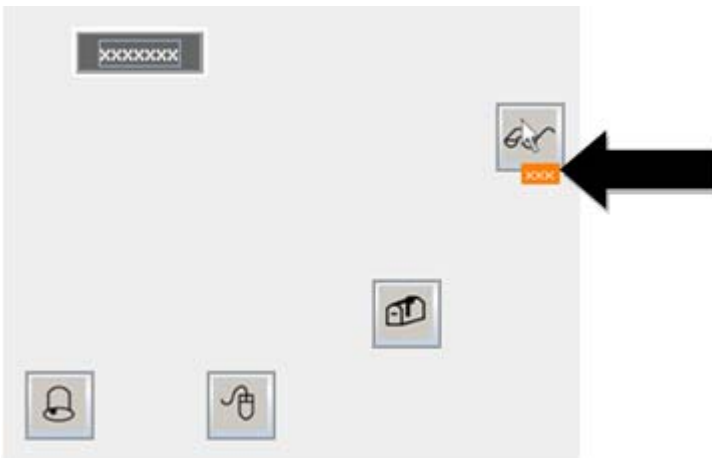
Figure 4: Tooltips were generated for each button as the user hovered over the button with the stylus.

Exercise three was given as the third application and involved a more complex visual

layout than exercises one and two where it was based on a simplification of the address

canvassing task used by the Census Bureau. In addition to the difference in software buttons,

the interface had a view of a map that changed in response to the selected software button

(Figure 1c). In this exercise, the target function was presented as a phrase that would show

up in black text on top of a white box that was located at the bottom of the screen (e.g.,

"Zoom in" and "Pan right" were included amongst the target functions). For a full list of all

target functions see Table 1). The 12 target functions in exercise three corresponded to 12

map interface buttons (Figure 3) that manipulated the map view (see Appendix A for a

presentation of all possible map views).

Table 1. Instructions given to user to indicate the target function.

Code	Target function	Button label (text)	Tooltip and instruction displayed to user
1	Pan right	Pan right	Pan right
2	Pan left	Pan left	Pan left
3	Pan up	Pan up	Pan up
4	Pan down	Pan down	Pan down
5	Zoom in	Zoom in	Zoom in
6	Zoom out	Zoom out	Zoom out
7	View assignment area	Assign. area	Show assignment area
8	Hide street names	Street names	Hide street names
9	Show user's location	You are here	Show your current location
10	Show address list	Address list	Show address list
11	View map spots	House loc.	Show address locations
12	Close map	Close map	Close map

In each exercise, the user was presented with 12 software buttons and a target

function over a series of 12 trials that were grouped in 16 blocks (a total of 192 trials). The

arrangement of buttons was held constant across the 12 trials in each block. Each of the 12

buttons was presented during a trial within a block. At the end of each block (set of 12 trials)

a dialog window appeared that provided feedback on the status of which block was just

accomplished (e.g., "End of block 1").

Location learning (or recall) was determined by administering a test at the end of each

exercise where subjects were presented with a blank screen that consisted only of the colored

rectangle they saw in the original exercise and one button with no label. The test involved

dragging and dropping the blank buttons that represented each of the target functions to a

location on a blank screen in attempt to place it where it was located in the former exercise.

During the test, the subjects were presented with a blank button that appeared above the target rectangle from the former exercise (see Figure 5a). Upon seeing the blank label and new target the subject tapped on the blank label to initiate an active move state (see Figure 5b). Once the button was active for moving, the subject dragged it to the location in which they recollected it being positioned from the former exercise (Figure 5c). Upon finding a position that they were satisfied with the user would tap again on the screen to "drop" the button. For each test, subjects would first perform one practice drag and drop move. Next they would complete a drag and drop move for each of the 12 targets. Each drag and drop was performed on an individual basis without being able to see their previous work.
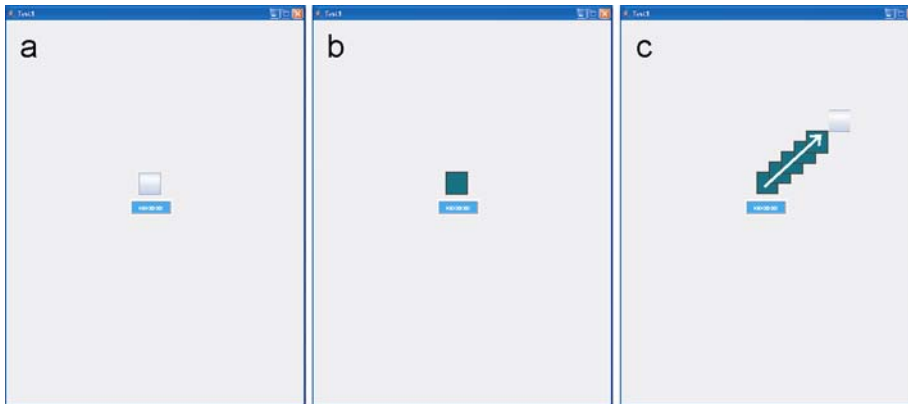
Figure 5: Example drag and drop in the location learning test.

Performance data included time to complete each trial, number of attempts to correctly select each button for each trial, and location learning test results (displacement distance measured in pixels; see Figure 6). Additional covariates captured by the software included data related to self-reported computer experience.
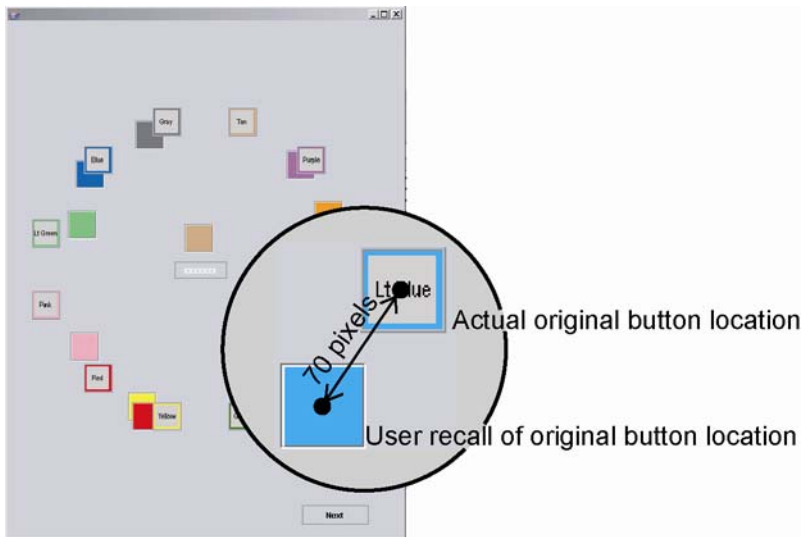
Figure 6: Results from a location learning test were measured in pixels.

3.3.4 Experiment design

The experiment involved a repeated measures design. Subjects in each age/sex group were randomly assigned to one of the four label treatments. To account for task experience, half of the subjects were randomly assigned to begin with exercise one and the other half began with exercise two. Each participant performed exercises on all three interface layouts.

3.3.5 Experiment procedure

Prior to performing the main exercises of the experiment, subjects were provided with a letter of informed consent (Appendix B). After agreeing to participate, each subject completed a color training task, a questionnaire that requested demographics and information about computer experience (Appendix C), three cognitive tests (Appendix D), and training on the primary exercises. A test script was used to ensure consistency (Appendix E). Two tablets were used in the study. One was set up for administering the background

questionnaire, the color training exercise, and the training for each primary exercise. The second tablet was used for administering the primary exercises.

The dimensions of the interface on all three exercises was 5 5/8 inches in width and 7 ½ inches in height. The computer used to display the interface was a Gateway Tablet PC M1300. This tablet had a 12.1-inch active matrix LCD color display (9 ¾ inches in width and 7 ¼ inches in height). The tablet was configured so that the display was in a portrait setting for the experiment.

The background questionnaire consisted of questions related to computer experience, education and employment status, age, gender, and an inquiry as to how the subject learned about the study. This questionnaire was developed as a three page form application and was administered on one of the two tablets.

Three cognitive tests were administered, including visualization, perceptual speed, and visual memory (administered in the former respective order). A description of each test is included below.

Visualization— Paper folding test. A common assessment used to measure visualization is the paper-folding test (Eliot & Smith 1983). The test includes a series of paper-folding problems. The subject sees a folded piece of paper that had one hole punched within it. Their task is to determine where the holes would be located in the paper if it were to be completely unfolded. This test involves 20 problems and is administered in two parts. The full test requires six minutes, that is three minutes per part (does not include the time it takes to provide instructions).

Perceptual speed— Number comparison test. In this test the subject evaluates two pairs of multi-digit numbers and indicates if the two numbers are the same or different. The

test consists of 96 problems which are administered in two parts (48 problems per part). The full test requires three minutes, that is 1 ½ minutes per part (does not include the time it takes to provide instructions).

Visual memory— Building memory test. In this test, the subject studies a map that includes a series of buildings. They are then tested on the specific location of each building. This test involves 24 test items and is administered in two parts. Four minutes were provided for study time and four minutes was provided for testing. Only the first part (12 test items) was administered. The full test for part one requires 8 minutes (not including the time it takes to provide instructions).

Once the cognitive tests were completed the overhead light was turned out to minimize glare on the tablet PC. The next step involved calibrating both tablets. Calibrating the tablets was performed by each subject to ensure that the tablet was sensitive to the users handedness and the way in which they used the stylus. At this point, the user was presented with a white screen where they tapped on four crosses that appeared at each corner of the screen.

The color training task familiarized the subjects with the 12 colors used in the first two exercises. The purpose of this task was to screen subjects for color blindness. It also served as a training exercise to familiarize the subjects with the color scheme that was used in the main exercise. In this training exercise, users were presented with a colored rectangle in the center of the application screen (Figure 7). Their task was to select the name of that color from a dropdown menu that was to the right of the colored rectangle. Once the user selected the name of the color they were instructed to tap the OK button below the dropdown menu.

If the user correctly selected the right button they would advance to the next color. If the user selected an incorrect color, a screen was displayed that read "Incorrect, the correct answer is blue" (for example). There were 12 possible selections that matched the 12 colors used in the primary exercises. Each color was shown three times for a total of 36 trials.
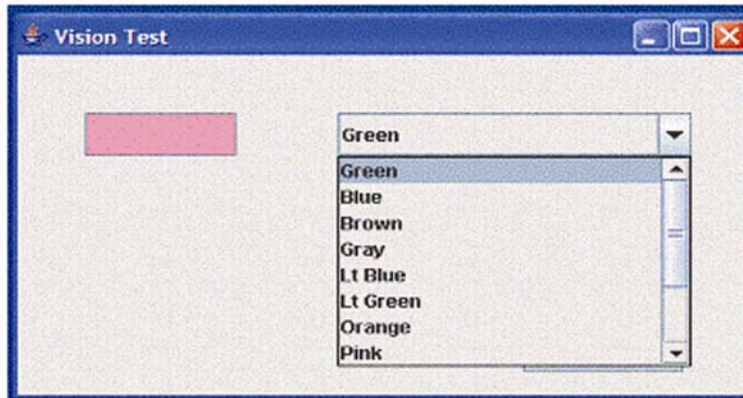
Figure 7: Color training interface showing target color and choice list for the name of the color.

Prior to each exercise, training was completed that involved practice on one full block of 12 trials. After the practice exercise the subject was told that they would work through 16 blocks and it was suggested that they should work as quickly and as accurately as possible. After completing each full exercise subjects were trained on how to perform the location learning test. The test was set up so that there was one practice trial. After the practice, the subject completed a drag-and-drop move for each of the 12 targets. After the test was completed the tablet displayed a message that instructed the participant to refer to the facilitator for instructions. After the subject completed all three exercises they were thanked for their time, provided with an incentive, and then excused to go. The experiment took

59

approximately one and a half to two hours to complete and each participant was compensated for their time with a $25 gift card to Target, HyVee, or the University Book Store.

3.3.6 Analysis

Analysis variables included total time in log seconds to complete each exercise, total number of attempts to find the correct button for each exercise, and the average distance in pixels of the user-specified location from the true button location based on the location learning test at the end of each exercise. Covariates included cognitive test scores for visualization, visual memory, and perceptual speed (Ekstrom et al., 1976).

Additional analysis variables were derived by fitting each subject's performance over blocks for an exercise to an exponential model that represented the learning curve for the exercise. The model took the form of $y= \alpha+\beta e^{-\gamma x}$, where y was total time for a block and x represented block sequence number (1-16). We derived three analysis variables from this model. The parameter α is the curve's asymptote, and represents performance time after the user is expert at the task. The value for $\alpha+\beta$ is the value of y where the curve intersects when x=0. This parameter function represents the performance time at the beginning of each exercise. The parameter γ reflects the steepness of decay (a decreasing function) in the time for each block as the user gains experience. This value represents the rate of learning as performance reaches its asymptote. Figure 8 illustrates each of these parameters in relationship to the exponential curve. This curve illustrates a steep decay which typically occurred where users had one of the uninformative label treatments. The other shape that was observed involved a flat slope. This shape typically occurred when users had one of the informative label treatments.
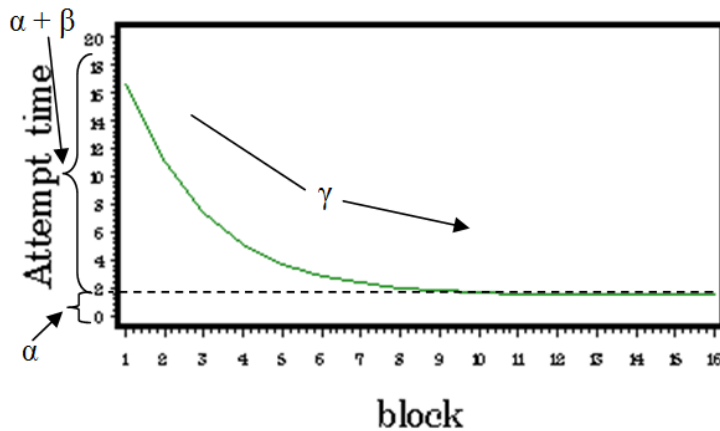
Figure 8: Example exponential curve.

Nonlinear least squares were used to estimate the learning curve parameters for each subject. Three bounds were set on the model so that estimates for α and β were greater than zero and the estimate for γ was greater than .01. Initially the general model ($y= \alpha+\beta e^{-\gamma x}$) was run for all subjects with starting points of α=2, β=10, and γ=.85. This first model converged for the majority of the observations (n=168), however failed for some (n=21). These instances occurred where there was a flat horizontal line or a very slow changing curve. Most of these shallow curves were associated with subjects who had either the informative graphic or text condition. Additional restrictions were placed on the model to fit curves for these subjects. The first alternative model involved setting α to zero to get the model $y=\beta e^{-\gamma x}$. This restriction provided a model that converged for 14 additional observations. The second alternative model involved setting both α and γ to zero to get the model $y=\beta$. This second restriction handled the remaining seven observations.

A mixed model analysis of covariance was performed that specified label treatments as between-subject treatments and layouts as within-subject treatments. The model included

classification factors of label (i.e., informative graphic, text, arbitrary graphic, no label),

layout (i.e., circle, grid, map), order of the first two exercises (i.e., circle first or grid first),

and their interactions was used to analyze the data. Continuous covariates in the model

included scores from the cognitive ability tests on visualization, visual memory, and

perceptual speed. Models were estimated for total time, total attempts, location learning, in

addition to the α, $\alpha+\beta$, and γ values derived from fitting the exponential learning curve

models. Least squares means were determined for each factor and its interactions with other

factors; pairwise comparisons among least squares means were made using the Tukey-

Kramer method. Slopes were estimated for the covariates, along with linear combinations of

slopes when an interaction between covariates and factors was significant.

3.4 Results

Table 2 presents the test results from the analysis of covariance for total log time,

total attempts, location learning (average distance), initial time ($\alpha+\beta$), experienced time (α),

and learning rate (γ). Because order was included in the model as a nuisance factor from the

experimental design, results for this effect are not discussed.

Table 2. Mixed effects table for total log time, initial and experienced time, learning rate, total attempts, and average distance.

Effect	Numerator Degrees of freedom	Denominator Degrees of Freedom[1]	Total Log Time F	p	Initial Time[2] F	p	Experienced Time[3] F	p	Learning Rate[4] F	p	Total Attempts F	p	Average Distance F	p
layout	2	86	1.12		0.58		0.41		0.41		1.67		0.46	
label	3	43	2.83	0.05	2.79	0.05	4.29	<.01	0.39		8.69	<.01	0.28	
layout*label	6	86	0.92		2.55	0.03	1.33		1.36		1.23		0.92	
order	1	43	5.32	0.03	0.00		2.68		0.15		6.11	0.02	5.30	0.03
layout*order	2	86	13.97	<.01	3.03	0.05	0.68		0.25		0.89		4.09	0.02
label*order	3	43	0.32		0.47		0.95		0.58		1.08		1.33	
layout*label*order	6	86	3.08	<.01	1.35		1.22		2.02		0.90		1.17	
visualization	1	43	2.71		9.39	<.01	0.16		1.03		3.13		9.41	<.01
visualization*layout	2	86	5.90	<.01	0.21		1.27		1.45		0.66		0.35	
visualization*label	3	43	1.25		2.20		1.29		1.02		0.99		0.80	
visualization*layout*label	6	86	0.85		0.71		1.26		0.67		2.00		0.18	
visual memory	1	43	8.99	<.01	2.88		1.83		4.50	0.04	27.39	<.01	0.25	
visual memory*layout	2	86	1.07		0.13		1.51		0.44		5.46	<.01	0.48	
visual memory*label	3	43	2.89	0.05	1.32		1.29		1.34		13.28	<.01	0.89	
visual memory*layout*label	6	86	0.69		0.30		1.61		0.97		1.95		1.46	
perceptual speed	1	43	11.90	<.01	4.64	0.04	8.79	<.01	0.01		0.85		1.47	
perceptual speed*layout	2	86	3.59	0.03	0.37		1.13		0.60		0.37		0.98	
perceptual speed*label	3	43	0.53		1.04		3.44	.03	0.48		0.26		0.35	
perceptual speed*layout*label	6	86	1.02		2.37	0.04	1.57		1.33		1.10		0.51	

1 Denominator degrees of freedom was not equal for all variables. One participant did not complete one of the three location learning tests. Therefore the denominator degrees of freedom are different for average distance than shown in the table.

2 Initial time represented α+β.

3 Experienced time represented α.

4 Learning rate represented γ.

Table 3. Means and standard errors of each label treatment for total log time, initial and experienced time, learning rate, total attempts, and average distance.

		Total Log Time	Initial Time	Experienced Time	Learning Rate	Total Attempts	Average Distance
Label	**Informative graphic**	5.70 (.06) A	1.52 (.10) A	1.03 (.13) A	.60 (.06) A	195.05 (2.41) A	63.18 (4.53) A
	Text	5.80 (.06) A	1.60 (.10) A	1.31 (.14) AB	.69 (.06) A	194.02 (2.46) A	66.19 (4.65) A
	Arbitrary graphic	6.31 (.10) B	2.82 (.16) B	1.54 (.22) AB	.76 (.10) A	229.07 (4.04) B	57.58 (7.50) A
	No label	6.48 (.07) B	2.86 (.11) B	1.77 (.15) B	.72 (.07) A	204.93 (2.76) C	57.62 (5.15) A

Table 4. Means and standard errors of each layout treatment for total time, initial and experienced time, learning rate, total attempts, and average distance.

		Total Log Time	Initial Time	Experienced Time	Learning Rate	Total Attempts	Average Distance
Layout	**Circle**	6.16 (.04)	2.41 (.09)	1.30 (.11)	.72 (.06)	207.73 (2.17)	66.18 (3.97)
	Grid	6.03 (.04)	2.07 (.09)	1.48 (.11)	.65 (.06)	203.60 (2.17)	59.60 (3.97)
	Map	6.04 (.04)	2.12 (.09)	1.46 (.11)	.71 (.06)	205.97 (2.17)	57.65 (3.98)

Label type had a significant effect on total performance time. Table 3 presents the least squares mean (LSM) and standard error (SE) for the time it took to complete each label treatment. Tasks performed using informative labels (i.e., informative graphic and text) required less time than uninformative labels (i.e., arbitrary graphic and no label). Table 4 presents the LSM and SE for the time it took to complete each layout treatment. There were no differences across layouts nor were there interactions with the different label types. Visualization, visual memory, and perceptual speed all had significant effects on total performance time. Table 5 presents slope estimates and differences for all significant cognitive ability effects.

Table 5. Slope estimates and estimated differences in slopes for total time for significant cognitive ability effects.

Effect	Total Log Time	
	Slope (SE)	p
visualization*layout		
Circle	-.0004 (.0113)	.97
Grid	-.0091 (.0113)	.42
Map	-.0366 (.0113)	.002
Circle – Grid	.0087 (.0010)	.43
(Average of Circle and Grid) – Map	.0319 (.0095)	.001
visual memory	-.0513 (.0171)	.005
visual memory*label		
Informative graphic	-.0026 (.0218)	.91
Text	-.0032 (.0353)	.93
Arbitrary graphic	-.1235 (.0452)	.009
No label	-.0760 (.0303)	.02
Informative[1] – Uninformative[2]	.0968 (.0342)	.007
perceptual speed	-.0122 (.0035)	.001
perceptual speed*layout		
Circle	-.0064 (.0043)	.14
Grid	-.0175 (.0043)	<.0001
Map	-0.0126 (0043)	0.004

1 Average of informative graphic and text
2 Average of arbitrary graphic and no label

For time to complete an exercise there was an interaction between layout and spatial visualization. Results in Table 5 indicate that spatial visualization had an effect on total time for the map layout, but not for the circle or grid layouts. For the map layout, the mean time to perform the task was shorter when subjects had a higher visualization score.

There was also a significant effect of visual memory on time to complete an exercise. In addition, there was a significant interaction between visual memory and label for time. There was no relationship for informative labels, but for uninformative labels (arbitrary graphic, no label), subjects with higher visual memory scores required fewer attempts to correctly identify the button (Table 5).

Finally, there was also a significant association between perceptual speed and time to complete an exercise for some layouts. Subjects with higher perceptual speed scores had shorter exercise times for the grid and map layout relative to subjects with lower perceptual speed scores (Table 5). No relationship was observed for the circle layout.

There was a significant effect of label treatment on the time it took at the beginning of each exercise ($\alpha+\beta$). There was also an interaction between layout and label for initial time to complete a block. For circle and grid layouts, there was a significant difference between informative labels (informative graphic and text) and uninformative labels (arbitrary graphic and no label) (Table 6). However for the map layout, there was more of a gradation between types of labels and this distinction was not as clear. The only significant difference in initial time for the map layout was between the text label and no label, with less time being required when text labels were given.

Table 6. Significant interactions of initial time (α+β) for the layout by label interaction.

Treatment	Mean (SE)				
Grid, Informative graphic	1.21 (.15)	A			
Circle, Informative graphic	1.39 (.15)	A	B		
Grid, Text	1.53 (.16)	A	B		
Map, Text	1.63 (.16)	A	B		
Circle, Text	1.65 (.16)	A	B		
Map, Informative graphic	1.98 (.15)		B	C	
Map, Arbitrary	2.25 (.25)		B	C	D
Map, No label	2.63 (.17)			C	D	E
Grid, Arbitrary	2.71 (.25)			C	D	E
Grid, No label	2.82 (.17)				D	E
Circle, No label	3.14 (.17)				D	E
Circle, Arbitrary	3.48 (.25)					E

There was a significant effect of visualization on initial time. Subjects with high visualization scores required less time to complete a block at the beginning of each exercise than subjects with low visualization (Table 7).

Perceptual speed was also related to initial block completion time, with an interaction between perceptual speed, layout, and label. Tests for differences in slopes presented in Table 7 indicated that the relationship between perceptual speed and initial block completion time was only significant for the circle layout without labels on the buttons and the map layout with informative graphic buttons. No other slopes were different from zero.

Table 7. Slope estimates and differences for initial time (α+β) for significant cognitive ability effects.

Effect	Initial Time	
	Slope (SE)	*p*
visualization	-.0481 (.0157)	.004
perceptual speed	-.0128 (.0059)	.04
perceptual speed*layout*label		
Circle, Informative graphic	.0149 (.0124)	
Circle Text	-.0079 (.0132)	
Circle, Arbitrary	-.0171 (.0212)	
Circle, No label	-.0656 (.0244)	0.008
Grid, Informative graphic	.00172 (.0124)	
Grid, Text	-.0213 (.0132)	
Grid, Arbitrary	-.0021 (.0212)	
Grid, No label	-.0180 (.0244)	
Map, Informative graphic	-.0342 (.0124)	0.007
Map, Text	-.0062 (.0132)	
Map, Arbitrary	.0193 (.0212)	
Map, No label	-.0167 (.0244)	

There was a significant effect of label treatment on the experienced block completion time (i.e., α), with informative graphic times less than those associated no label (Table 3). There was a significant effect of perceptual speed on experienced time that was mediated by label type (Table 8). In general, higher perceptual speed scores were associated with lower experienced block completion times, but uninformative labels had a significantly steeper (negative) slope than informative labels. Additionally, the no label treatment had a significantly (negative) steeper slope than the arbitrary graphic label.

Table 8. Slope estimates and differences for experienced time (α) for significant cognitive ability effects.

	Experienced Time	
Effect	*Slope (SE)*	*p*
perceptual speed	-.0243 (.0082)	.005
perceptual speed*label		
Informative[1] – Uninformative[2]	.0435 (.0164)	.01
Informative graphic – Text	-.0078 (.0160)	0.63
Arbitrary – No label	.0592 (.0286)	.04

1 Average of informative graphic and text
2 Average of arbitrary graphic and no label

The learning (or decay) rate (γ) from individual block completion time curves was related to visual memory scores (slope estimate= .0366, SE=.0173, *p*=.04). Subjects with high visual memory had a faster learning rate than subjects with low visual memory. Interpreting this variable was difficult because the decay rate in the curve was related to the initial performance time. That is, subjects who did poorly in the initial blocks all significantly improved their performance, although the rate of learning varied somewhat. In contrast, subjects who performed well initially did not improve substantively (near zero decay rate). In essence, the measure was flawed because the decay rate for good initial performers, who did not need to learn, were similar to poor initial performers who did not improve their performance.

Label type had a significant effect on total attempts to complete an exercise (Table 2). Tasks performed using informative labels required fewer attempts than uninformative labels and tasks performed using the no label treatment required fewer attempts than the arbitrary graphic treatment (Table 3).

There was a significant relationship between visual memory on attempts, with an interaction between visual memory and layout. In general, subjects with high visual memory

required fewer attempts than subjects with low visual memory (Table 9). However, visual

memory had a stronger effect on total attempts for the circle layout (i.e., a steeper slope) than

it did for the grid and map layouts.

There was also a significant interaction between visual memory and label on

attempts. Visual memory had no impact on total attempts for informative labels and for no

label, but a higher visual memory score was associated with fewer attempts for the arbitrary

graphic label.

Table 9. Slope estimates and differences for total attempts for significant cognitive ability effects.

Effect	Total Attempts	
	Slope (SE)	*p*
visual memory	-3.7358 (.7138)	<.0001
visual memory*layout		
Circle	-6.2152 (1.0364)	<.0001
Grid	-2.6123 (1.0364)	.01
Map	-2.3797 (1.0364)	.02
Grid – Map	-.2326 (1.3015)	.86
(Average of Grid and Map) – Circle	-3.7192 (1.1271)	.001
visual memory*label		
Informative graphic	-.1607 (.9108)	.86
Text	-.6885 (1.4712)	.64
Arbitrary graphic	-13.0686 (1.8864)	<.0001
No label	-1.0253 (1.2650)	.42
Informative[1] – Uninformative[2]	6.6224 (1.4276)	<.0001
Arbitrary – No label	-12.0432 (2.2713)	<.0001

1 Average of informative graphic and text
2 Average of arbitrary graphic and no label

There was a significant relationship between the visualization score average distance

between the subject's recalled button location and its true location (slope estimate= -2.2419,

SE=.7307, *p*=.004). Higher visualization scores were associated with smaller distances,

which were indicative of higher location recall accuracy.

3.5 Discussion

Our primary interest was in determining how spatial ability, as measured through spatial visualization, might affect user performance for simple button recall tasks. For the time it took to complete an exercise, visualization was only important when labels were uninformative or when the more complex map interface was being presented.   This finding indicated that performance differences for subjects of low and high spatial visualization ability were smaller for a familiar color-matching task and the map interface which supported an atypical task, supporting one of our initial hypotheses. These differences could have occurred as a result of subjects with low spatial visualization having problems developing new declarative memories and survey knowledge for unfamiliar labels. All users had a pre-existing declarative memory for the color-matching task performed in the circle and grid exercises. This was evident from observing each user perform the color training task which was administered at the beginning of the experiment. Because colors are well known to most people it did not place a high demand on working memory to encode new memories for color representations. In contrast, the map task involved operations that were not common knowledge (e.g., view assignment area and view map spots). As a result, we think the map layout was harder to learn for subjects with low spatial visualization because it placed cognitive demands on them to encode new memories for labels which they did not have pre-existing declarative memories for.

As seen with Shepard and Feng's (1972) work, success on the paper-folding task is highly contingent upon multiple abilities. The punched holes test for spatial visualization by Ekstrom et al. (1976) has the potential to be related to performance in three ways. First, there

the task involves several serial mental operations (individual folds) to arrive at an answer. Processing this information is highly dependent upon short-term memory capacity. Second, a person's capacity to effectively ignore other perceptual information surrounding the current problem will help them to avoid distractions that would interrupt their mental process. This skill is beneficial to successfully performing visual searches to find objects in interfaces. Third, the ability to effectively maintain a visual memory of the position of the punched hole at the same time as performing a mental rotation is relevant. This ability is an indicator of visual memory and dual-tasking ability. These skills are beneficial to remembering button locations on interfaces and working through problems that involve multiple layers.

On average, subjects with high visual memory required less time to complete an exercise than subjects with low visual memory. Further, subjects with higher visual memory scores tended to take less time to complete an exercise with uninformative labels than subjects with low visual memory. These findings suggest that subjects with a higher capacity for visual memory stored a mental image of the interface layout more quickly than subjects with low visual memory. Craik and Lockhart (1972) offered a model for level of processing in which deeper encoding results in better memory. For the easier informative and text labels, it is possible that deep processing did not occur because subjects relied on pop-out information as opposed to encoding location. In contrast, elaborative encoding occurred for uninformative labels (i.e., an encoding of location and/or a color association with a graphic). Further, it is possible that the encoding that occurred for subjects with high visual memory may have been deeper than for subjects with low visual memory when labels were not informative. These findings are consistent with our second hypothesis that performance

differences for subjects of low and high spatial ability would be smaller with informative button labels than with uninformative labels.

Perceptual speed was also an important factor for total time depending on the layout used. Perceptual speed is directly related to spatial ability (Lohman, 1996) in that it taps the ability to recognize perceptual contrast (i.e., differences between like objects). Subjects with high perceptual speed took less time than subjects with low perceptual speed when performing the color task with the grid and map interfaces. It is possible that subjects with lower perceptual speed scores struggled to conduct their visual search and make perceptual contrasts using the grid structured layouts. Like visual memory, perceptual speed is an ability that has been found to be related to spatial ability, and thus these results are consistent with our hypothesis that subjects of high spatial ability would have better performance time than subjects with low spatial ability.

Performance at the beginning of an exercise may be more sensitive to layout and labeling than after the subject has had time to learn how a software interface works. We found that initial block performance times for the simple color match layouts were clearly better with informative labels than when uninformative labels were used. An interesting deviation was observed with the map interface. Here the distinction in performance among the different label types was less clear, with only text labels requiring significantly less time for the initial block than for buttons with no labels. The differences between the color match and map-based interfaces may help explain why the distinction between informative and uninformative labels was not as strong for the map exercise. The informative labels in the color-matching exercise involved matching a color to a color. In contrast, the informative labels in the map exercise involved matching a non-conventional graphic or text phrase to a

function that the subject may not have been familiar with. That is, the lack of familiarity with basic map functions and associated informative text and graphical labels likely presented a cognitively more demanding task that mitigated the performance gains observed for the simpler and more familiar color match task and informative buttons. This finding is consistent with other research that has reported that cognitive abilities are most significant when a complex task is involved (Pak et al., 2006).

The main effect of visualization on initial performance time is consistent with many of the findings in existing literature that have reported a relationship between spatial ability and computer usage. The results supported our hypothesis that that in general subjects with high spatial visualization would tend to require less time to perform a visually oriented task in the beginning than those subjects with low spatial visualization.

The significant effect of label on final performance time indicated that label effects were persistent. The difference between the informative graphic and no label treatment remained throughout the exercise even into the time required to complete final blocks of the exercise. These findings suggest that either learning was still occurring at the end of the experiment for the no label condition or that a persistently poorer performance time could be expected for an interface that has buttons without labels. The finding that there was no difference between the no label treatment and the arbitrary label treatment lends support that no label could be just as good as a bad label.

Perceptual speed was correlated with performance time for the experienced performance time. Subjects with high perceptual speed took less time at the end of each exercise than subjects with low perceptual speed. Specifically, these findings were observed to a greater extent with uninformative labels. There was even more of an effect when the no

label treatment was involved. It is possible that visual search times may have been longer for subjects who had lower perceptual speed scores or that subjects with lower perceptual speed had a harder time ignoring distracters (e.g., labels that were similar to one another such as green and light green).

Visual memory was an important determinant in the number of attempts a user took to find the correct button when conducting an exercise. Uninformative labels required more attempts than informative labels for subjects with low visual memory. In addition, subjects with low visual memory required more attempts than subjects with high visual memory using the arbitrary label treatment than the no label treatment. These findings provide additional support to the hypothesis that performance differences that relate to the labeling of the button will be greater for subjects of lower spatial ability relative to those with higher abilities.

The association between visual memory and number of attempts was also mediated by the layout of the interface. In particular, the circle layout on average required more attempts than the grid and map layouts for subjects with low visual memory relative to subjects with high visual memory. In this case, performance differences in low and high visual memory users may be related to the familiarity and efficiency of the button layout. The grid and map labels were organized in two linear rows and thus common to the type of structure used in frequently used applications such as Word and Excel. In contrast, the circle layout involved more space for visual search and did not provide a familiar linear arrangement that was useful for ordering.

The significant association of visualization scores with the accuracy of button location recall (i.e., average distance from the true button location) indicated that subjects with low spatial ability had a poorer recall for spatial locations than high spatial ability

subjects. It is generally accepted that most people create mental maps of spatial configurations to keep track of the comprehensive layout of space. Creating a mental map is similar to the ability to encode spatial knowledge in a survey (i.e., map-based or configuration) framework. Prior to developing survey knowledge, a person establishes declarative and procedural knowledge, which involves collecting and encoding list information about landmarks and routes, respectively. The ability to develop survey knowledge is positively correlated with spatial visualization scores. In this experiment, knowing where each button was located in the set of 12 buttons would be like developing survey knowledge. Thus, a subject with high visualization score might be expected to more rapidly learn the locations of the software buttons.

CHAPTER 4. THE ROLE OF SPATIAL ABILITY IN
USING MAP-BASED GUIDED APPLICATIONS

Abstract. Spatial ability is often cited as the individual cognitive ability most likely to affect software user performance (Dahlbäck, Höök, & Sjölinder, 1996; Egan, 1988; Vicente, Hayes, & Williges, 1987; Egan & Gomez, 1985). There remain questions about how the component factors involved, such as visualization and orientation, in an individual's spatial ability influence how software is used. We conducted a study to examine how these factors affect user performance for map software via a simulated field task to verify address locations on a map. Twenty-four subjects were randomly assigned to use software that offered verbal guidance to assist the user or to use the same software without any guidance. We hypothesized that subjects with low spatial ability would require more time, be less accurate, and be less efficient with their button usage while working with complex map-based graphical user interfaces relative to subjects with high spatial ability. Additionally, we hypothesized that extra screen information on the interface might help to improve task performance of subjects with lower spatial ability, more so than improving performance of subjects with high spatial abilities. Findings indicated that there was a relationship between user performance and spatial ability for time, accuracy, and pan button usage. Furthermore, these findings supported the hypotheses that spatial visualization and orientation are dissociable skills for accuracy and pan button usage. We did not find evidence to support benefits or disadvantages of using guided systems.

4.1 Introduction

Human computer interaction (HCI) researchers have recognized the importance of individual differences in cognitive abilities for designing effective software applications. Egan (1988) found that user performance in accessing information via software interfaces is affected by an individual's spatial ability, memory, reasoning ability, verbal aptitude, and possibly personality. Of these abilities, spatial ability has been found to be one of the strongest predictors of human computer performance (Dahlbäck, Höök, & Sjölinder, 1996; Egan, 1988; Vicente, Hayes, & Williges, 1987; Egan & Gomez, 1985). Lohman (1996) defined spatial ability as "the ability to generate, retain, retrieve, and transform well-structured visual images" (p. 98). The relationship between user performance and spatial ability is believed to arise from the fact that interacting with an interface is by nature spatial.

Spatial visualization and spatial orientation are two sub-factors of spatial ability that have been found to be correlated with a software user's performance. Spatial visualization has been defined as the "ability to manipulate or transform the image of spatial patterns into other arrangements" (Ekstrom, French, Harman, & Dermen, 1976, p. 173). Spatial orientation has been defined as the "ability to perceive spatial patterns or to maintain orientation with respect to objects in space" (Ekstrom et al., 1976, p. 149). Furthermore orientation has been described in two forms (Kozhevnikov & Hegarty, 2001). One form involves the mental rotation of objects. The other form is dependent on a person's frame of reference in relationship with the real world. Visualization and orientation have been shown to be distinct from one another yet most research on computer performance examines either one or the other (Pak, Rogers, & Fisk, 2006).

Pak et al. (2006) considered multiple sub-factors of spatial ability in a study that investigated an information search task. They found that spatial orientation ability was related to performance with tasks that were high in their navigation requirement. In this case, the cube comparison test was used to assess orientation ability. Cube comparison involves skills that tap a person's ability to make allocentric configurations (i.e., solving problems that are independent of using sensory feedback to determine direction). They did not discover any significant results relating to visualization. Their explanation for these findings was that a relationship with orientation ability is dependent upon the nature of the task. In their case, because navigation played an important role in completing the task, orientation ability was the best predictor. While this group made progress in testing orientation and visualization, their assessment of the mental rotation aspect of orientation did not involve evaluation of the person's frame of reference (i.e., the use of egocentric configuration that involves re-mapping of their own current physical position).

The importance of spatial ability in software user performance suggests that it may be advantageous to the user if software systems were able to accommodate individual differences in spatial ability. Zhang and Norman (1994) argue that when considering a cognitive task, it is never solely dependent upon the internal mindset of the users nor is it solely external to effective design of interface displays. Both individual differences and system design have the potential to influence computer performance. Büring, Gerken, and Reiterer (2006) found that persons with low spatial ability showed better performance in using a scatterplot-like interface with a real-time overview feedback window relative to an interface without the overview window. The feedback consisted of a comprehensive view of the whole area with an indicator "you are here" box which provided reference to where the

user was at when they were zoomed in at magnified levels. In contrast, high spatial ability users were hindered by the system with the overview window. Büring, et al. (2006) provided some insight into the influence of individual systems. Unlike modern forms of guided applications that provide "next step" cues, their study only provided map orientation cues. We wanted to explore the potential benefit of a guided system (i.e., inclusive of cues specific to what actions should be taken next) to users with varying levels of spatial ability.

To investigate these ideas, we collaborated with the United States Census Bureau in developing a study that focused on the task of address verification. In this task, field staff evaluate whether the location of an address on the ground is properly represented as a small dot on the map. If map error is detected, the location of the housing unit is modified to correct the error. This task can be cognitively demanding and involves both mental rotation and orientation. As the field staff member identifies a housing unit, they must correctly find that position on the map before going to the door. In this case, they make a comparison between a map and the real-world, which may also involve mental rotation of the map to reflect the orientation of the housing unit on the street.

Our goal was to investigate whether guided applications will help to bridge the gap between persons with low and high spatial ability by reducing the need for spatial visualization. Sein, Olfman, Bostrom, and Davis (1993) found that with appropriate training and interface design the gap between high and low visualization users could be reduced and even reversed. In our study, users were asked to identify whether housing units were present on the ground situation (represented in photographs) and whether they were accurately located on the map. Subjects were randomly assigned to two interfaces, one that offered guidance on the protocol and another that was exactly the same, but did not present guidance.

For the interface treatment that did not provide protocol guidance (unguided treatment), we hypothesized that subjects with low spatial ability would require more time, be less accurate, and require more tool usage interpreting map content and working with complex graphical user interfaces relative to subjects with high spatial ability. We expected that guided interface would help reduce cognitive demands for low spatial users by offsetting the demands on memory to recall the protocol sequence and allowing low spatial ability users to devote more effort toward the complicated spatial transformations. In contrast, because a subject of high spatial ability is facile at storing and manipulating spatial information, we expected that such an individual to readily develop efficient methods of using map software to perform tasks. In addition, we expected that extra screen information on the interface would be less helpful to this group. In addition, we hypothesized that spatial visualization and spatial orientation would be dissociable for performance measures such as time to perform the task and accuracy.

The following sections include a review of literature specific to the relevant aspects of individual differences, spatial ability, and working memory associated with the address verification task followed by a detailed account of an experiment that was conducted to explore the effective use of guided versus unguided interfaces.

4.2 Background

4.2.1 Individual differences

Dillon and Watson (1996) and Benyon, Crerar and Wilkinsin (2001) provided a thorough review of the major types of individual differences acknowledged in the current

literature. Among these differences they presented categories of psychological, physiological, and socio-cultural characteristics (for a full summary, see Chapter 2). Psychological differences include characteristics specific to cognitive abilities. Cognitive abilities and aptitudes are often defined as a person's capability to perform a variety of mental functions such as perception, memory, and reasoning. Additionally, these abilities are often associated with larger constructs such as intelligence. Spatial ability is an ability often expressed in these terms. These authors noted that individual differences related to psychological characteristics are most effective when considering interface design because they provide specific user characteristics that can be used to influence software design. In contrast, physiological and socio-cultural differences only offer insight into user behavior at a broad level (e.g., gender, physical size, language, etc.) that is difficult to incorporate into interface design.

Another component of individual differences is the amount of knowledge or experience a user has. Kalyuga, Chandler, and Sweller (1998) conducted research investigating individual differences in expertise and the relationship with interface design. Their main finding was that novice users preferred graphics (i.e., maps and diagrams) that were integrated with text to explain the image, whereas experts were more efficient with graphical diagrams that did not include the extra text. Their theory was that as the person's expertise changes, so do their needs from the interface. Since the mental process of integration is cognitively demanding, adding extra information to the stimulus was found to help novice users understand the material. In contrast, expert users experienced unnecessary cognitive load in processing the extra information because it was found to be redundant. In this case, an interface that presented graphics alone (without the supplementary text) would

have been most suitable to their expert users. This research illustrated how individual differences can affect user performance when the amount of experience is taken into account. Expert users are expected to develop more advanced strategies (i.e., probably more efficiently) than novice users. The individual differences considered here would be classified as flexible because novice users could become expert users over time. Thus the limitation on this type of individual difference is that it is potentially highly vulnerable to change.

4.2.2 Spatial ability

Spatial ability is consistently linked with computer performance (Dahlbäck, Höök, & Sjölinder, 1996; Egan, 1988; Egan & Gomez, 1985; Vicente, Hayes, & Williges, 1987). Lohman's (1996) proposed that spatial ability involves the capacity to perform multiple operations on visual images (i.e., generating, retaining, retrieving, and transforming). This definition implies that spatial ability is not a unitary construct. Specifically, there are multiple dimensions that deal with context, scale, and variability of the degree to which serial operations are involved.

Spatial ability may vary with age and gender. Many studies investigating demographic associations with spatial ability have found that as an individual gets older, spatial ability can decline. For tasks that involve mental rotation, speed may increase but errors also increase (e.g., Clarkson-Smith & Halpern, 1983). Many studies have indicated that spatial ability tends to be higher for men relative to women for mental rotation (e.g., Bryant, 1982) and for spatial perception (understanding the relationship of an object in relationship with other surrounding objects when distracting information is present) (Linn & Peterson, 1985).

Another important component for the interpretation of map content is the subset *orientation* construct of egocentric transformations. Spatial orientation has been defined as the "ability to perceive spatial patterns or to maintain orientation with respect to objects in space" (Ekstrom et al., 1976, p. 149). Orientation is related to the frame of reference, specifically the egocentric and allocentric (object-based) perspectives. Burgess (2006) suggested that these perspective taking abilities exist in parallel. Egocentric perspective involves the ability to imagine a reoriented-self (perspective taking). Allocentric perspective involves the ability to imagine the rotation of objects from a fixed perspective (mental rotation). Allocentric perspective is used when performing tasks that involve visualization skills. Visualization is a form of spatial ability that is also of interest to this work and will be further addressed in a later section. The following sections provide details on differences related to egocentric and allocentric transformations, the influence of scale, example subset constructs of each perspective, and a presentation of common techniques used to assess these skills.

4.2.2.1 Egocentric perspective

Egocentric perspective involves sensory feedback that contributes to the sense of direction of oneself in physical space. This type of configuration is often involved with large scale interpretation. An example is reading a map to find one's way in a large city. This ability is comparable to the skill that would be required to make a judgment about direction when using a map compared to the real-world. Hegarty, Richardson, Montello, Lovelace, and Subbiah (2002) and Hegarty, Montello, Richardson, Ishikawa, and Lovelace (2006) found a significant difference in types of spatial ability that exist for large-scale (for example,

84

the information that a person receives about orientation when standing in an open space) and small-scale applications (for example, the spatial skills one would use to solve a rubik's cube). They found that large-scale spatial processing skills required sensory feedback. Thus large-scale spatial ability should be measured when conducting an outdoor field study to investigate spatial ability.

Shepard and Metzler (1971) found that reaction time increases as the angular difference of orientation increases. Furthermore, people are faster and more accurate in naming and pointing at objects in front and behind them than those to the left or right when imagining perspectives (Bryant & Tversky, 1999; Hintzman, O'Dell, & Arndt, 1981). Kozhevnikov and Hegarty (2001) also found results that supported these ideas on pointing accuracy where absolute angular error increased with the angular deviation of one's imagined heading from the orientation of the array.

A common technique used to measure spatial ability involves the use of psychometric tests. Hegarty et al. (2002) suggests that self-reported measures are a more promising approach to measuring large-scale environmental spatial ability than psychometric testing. Other techniques for measurement include sketch maps and perspective taking tests. Kozhevnikov, Motes, Rasch, and Blajenkova (2006) created a test known as the Perspective Taking Assessment to examine this ability. In this test the user has to: 1) determine the position of a target stimulus in relationship to a given perspective that varied in direction and 2) store a memory of the angle between the perspective and the target object and then 3) map that position onto an array of arrows.

4.2.2.2 Allocentric perspective

In contrast to egocentric configurations, allocentric configurations involve "figural space" manipulations. Figural space has been defined as small in scale relative to the body and external to the individual (Hegarty et al., 2006). An example is solving the 2-dimensional problems presented in the game Tetris. In this case, playing tetris involves responding to small scale visual cues within an isolated display. Physical motion is not involved because the game is typically played sitting down. Therefore internal resources (e.g., vestibular senses) that provide feedback for orientation are not used and thus the task would be considered external to the individual. Lohman (1996) and Carroll (1993) suggested that there are five constructs of spatial ability that include visualization, speeded rotation, closure speed, closure flexibility, and perceptual speed. These constructs involve allocentric configurations because each require cognitive skills specific to solving transformation problems that are small in scale and thus do not have large requirements for sensory feedback. For example, a user study within a controlled lab would involve small-scale processing because the person would be isolated to a small space and the confines of a computer display. Therefore small-scale spatial ability would be the primary measurement of interest. Psychometric testing is often used as a paper-based measurement tool to assess spatial ability in user studies. A popular form of these tests is found in Ekstrom's Manual for Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Dermen, 1976). This kit includes a diverse battery of cognitive tests that covers the constructs of spatial ability presented by Lohman (1996) and Carroll (1993).

Visualization is an aspect of spatial cognition that involves "ability in manipulating visual patterns, as indicated by level of difficulty and complexity in visual stimulus material

that can be handled successfully, without regard to the speed of task solution (Carroll, 1993, pp. 362-363)." A common assessment used to measure visualization is the paper-folding test (Eliot & Smith, 1983). In this test the user must use cognitive processes to do two things: 1) make a multi-step transformation of the folded piece of paper and 2) store a visual memory of the holes in short-term memory.

Sein, Olfman, Bostrom, and Davis (1993) investigated visualization ability in relationship to usage of three applications (email, modeling software, and operating systems). They found that persons with high visualization skills learned fastest on all of the applications. They stress the importance of direct manipulation interfaces suggesting that they help the user to visualize system activities. In doing so, Sein et al. suggests that it alleviates the onus on the user to form a detailed internal representation. Thus the low visualization users can perform just as well as high visual users. Kozhevnikov and Hegarty (2001) found that visualization is dissociable from measures of orientation in perspective taking. One distinct difference is that visualization involves the need to remember a sequence of steps. In contrast, the perspective taking test simply involves the memory storage required to do a mental rotation.

Additionally, individual differences in visualization involve visual memory which may affect the ability to remember menu structures. Other research has shown that this ability carries over to how users interact with software interfaces (Ziefle & Bay, 2006). Just as one would expect visual memory to come into play in navigating physical spaces in order to remember landmarks, we expected that visual memory will also influence software usage. Consider the Büring et al. (2006) study that compared differences in using a zoomable interface with and without an overview window. The interface without the overview window

required users to remember where they were at in the current view. In contrast, the interface with the overview relieved the user of the burden on memory to remember where they were at in the comprehensive picture.

Logical reasoning was described by Carroll (1994) as involving retrieval of both meanings and algorithms from long-term memory that is followed by a performance of serial operations on the materials retrieved. Allen (1994) investigated the relationship between perceptual speed and logical reasoning with performance using an information retrieval system that presented information in two ways (rank-ordered by relevance and non-rank-ordered). There was no effect of perceptual speed, however there was an interaction between logical reasoning and interface type. Allen concluded that users with low reasoning skills performed better using a rank-ordered system than a non-rank-ordered system.

While reasoning is not listed as one of the five constructs acknowledged by Lohman (1996) and Carroll (1993), it is widely accepted that the capacity to reason is related to spatial ability. Researchers have suggested that working memory capacity involves the ability to manipulate information and suppress distracting information (Engle, 2002). Furthermore, other researchers have supported this idea and found working memory capacity to be related to reasoning (Ackerman, Beier, & Boyle, 2002). These abilities are important to spatial transformations because they help a person to control their attention and develop effective strategies to carry out complex mental rotations. More information specific to working memory is presented in the next section.

4.2.3 The influence of attention on spatial alignment

When addressing working memory it is helpful to have an understanding of the influence of attention. A review of the cognitive science literature reveals that there are several theories about how attention should be defined. A frequently quoted definition comes from classic work by William James, in his *Principles of Psychology* (1890). He reported that attention "is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought" (p. 403-404).

Feature integration theory (FIT) proposes attention in the form of a spotlight (Wolfe, Cave, & Franzel, 1989) with limited capacity that processes information in a serial manner (Treisman & Gelade; 1980; Treisman, 1986). The spotlight metaphor suggests that attention can only be focused on one item or a group of items at a time and can only activate awareness of a certain amount of space. The car wash has been a popular analogy of this theory. For example, picture cars going through a car wash one at a time. It is only possible to wash (i.e., process) one car at a time. This model of attention suggested that slope is determined by plotting reaction times as a function of the number of distracters. Distracters involve irrelevant information surrounding the target object. As the number of distracters increases reaction time required prior to recognizing the target object also increases.

The FIT model proposed two ways by which we process visual information: pre-attentive and attentive processing. Pre-attentive processing allows for quick feature searches that aren't affected by number of distracters. This form of visual processing is limited to a small set of feature differences such as color, size, motion, or orientation. For example, pre-attentive processing allows one to quickly spot a red item among several green items regardless of the amount of green items. Reaction time does not change in relationship to the

number of distracters because there is only one feature difference that is required for object recognition. In contrast, attentive processing is required for conjunction searches. Conjunction searches are performed when visual information must be bound for object recognition. For example, an object with disconnected joints (not conjoined) would be more challenging to find in a set of other objects with disconnected joins by comparison to a solid (conjoined) object.

These concepts are extendable to understanding spatial alignment where they provide insight into how and when a person is choosing to make a mental transformation. For example, consider the use of a map on a handheld device being used to get information about a specific location. In this scenario one has multiple options available. They might choose to zoom in to the general vicinity through a pre-attentive selection and then use attentive processing once zoomed in to make the necessary alignment between streets on the map to streets in their immediate surroundings. This strategy would be beneficial to persons with a low capacity to control their attention because many of the non-relevant map features present at a zoomed out level could potentially serve as distracters. A second strategy would involve performing an attentive selection even at the most zoomed out level. Then upon zooming in they already have the transformation complete and thus are most efficient with respect to performance.

4.2.4 The influence of working memory on dissociating between visualization and orientation

Similar to spatial ability, it is commonly accepted that working memory is not unitary (Baddeley & Hitch, 1974). That is, as opposed to there being one type of memory there are

multiple memory system types. Tulving (1985) used experiments to test cognitive processes occurring in multiple memory systems. He found that a particular lesion or type of stimulation that affects the performance on one task but not on another is suggestive of the existence of multiple memory systems. Tulving's research investigated performance for two tasks, word recognition and fragment completion. Word recognition involves presenting a subject with a list of words to remember and then presenting them with a different list after memorization in which they are tested to see how many of the words they can successfully recognize. Word fragment completion involves presenting a subject with words that are not fully completed (e.g., "h_gh_ay"). The subject is tested to see how many partial words s/he can complete correctly. A comparison of these tasks would suggest that they should involve different skills and memory retrieval. Tulving identified areas where performance for the two tasks performed by the same person was uncorrelated and suggested the results provided strong support for the idea that cognitive processes occur in multiple memory systems. Tulving's (1985) research can be applied to examine the differences between spatial visualization and orientation (i.e., perspective taking). Much of the current literature supports the notion that these two skills are dissociable from one another (Pak et al, 2006). If there is a lack of correlation among skills and performance measures within a model, it is likely that different memory systems are in play.

4.2.5 Guided versus unguided applications

One area of opportunity that exists for studying tailored interfaces to suit various levels of spatial ability is the study of guided versus unguided applications. Researchers have commonly found that when designing computerized systems for students learning by

discovery is generally observed to provide greater gains for students with high cognitive abilities and greater losses for students with low abilities (Berger, Lu, Belzer, & Voss, 1994). Learning by discovery (i.e., involving contemplation and reflection) is often the result of unguided systems which can be beneficial to persons with high abilities where they will lead to a strong memory of the interface. The stronger memories are the result of a deeper level of encoding. In contrast, guided systems often are more useful than unguided systems for persons with low ability. Zhang (1997) has found that the externalization of help aids and extra information (i.e., like those found with guided systems) can be advantageous if the benefit of using such aids can offset the costs associated with using them (e.g., the risk of excess clutter that may result in an increase on the demands of cognitive processing). The limitation on much of the work that has been conducted on guided versus unguided applications is that few have investigated the role of individual differences.

4.3 Method

4.3.1 Overview

We conducted a study to evaluate whether spatial ability affected the ability of users to perform an address matching task on two different interfaces that varied in the area of provided guidance. We recruited 24 subjects to perform 10 test scenarios. Each subject was randomly assigned (according to spatial ability that was determined in the first session) to a specific interface that was guided or unguided. Software was developed to guide the subjects through a multi-step procedure that was specific to each scenario. The application recorded time to perform each step in the procedure, number of attempts to match each address, number of attempts to fix the map, accuracy in fixing the map, and overall tool usage.

There were two hypotheses specific to time. First, we expected that subjects with high spatial ability (i.e., visualization and/or perspective taking ability) would have better performance time using the unguided interface relative to subjects with low spatial ability. Second, we hypothesized that subjects with low spatial ability (i.e., visualization and/or perspective taking ability) would have similar or better performance time using the guided interface relative to subjects with high spatial ability. For accuracy in placing map spots on the map, we hypothesized that subjects with high spatial ability would have better map spot accuracy when adding spots relative to subjects with low spatial ability. For tool usage, we hypothesized that subjects with high spatial ability would have higher reset map, zoom, and pan efficiency (i.e., less usage) relative to subjects with low spatial ability.

4.3.2 Participants

The 24 subjects were recruited through fliers posted on the Iowa State campus, local grocery stores, coffee shops, the public library, and word of mouth. This method was used because it mimicked recruiting strategies used by the Census Bureau to recruit address listing staff. Prior to advertising, four categories were specified which were used for recruiting subjects. These categories were: 1) female younger than 30 years of age, 2) male younger than 30 years of age, 3) female older than 30 years of age, and 4) male older than 30 years of age. Upon first contact, subjects were classified by gender and age category. Our goal was to include six subjects within each category. The final demographic distribution included six females that were 18-29 years of age, six males that were 18-29 years of age, three females that were 30-39 years of age, one male that was 30-39 years of age, one female that was 40-

49 years of age, two males that were 40-49 years of age, two females that were 50-59 years of age, two males that were 50-59 years of age, and one male that was 60+ years of age.

Additional demographic and computer experience information was collected at the time of the study. Out of all 24 subjects, eight were undergraduate students (33%), three were graduate students (13%), and thirteen were not students (54%). Fourteen were employed full-time (58%) and eight part-time (33%), one subject was neither employed nor retired (4%), and one subject was retired (4%). The majority of subjects had a fair amount of experience working with computers. Twenty subjects (83%) reported having a high level of experience using a personal PC or laptop computer. Many also had experience using common software applications where 23 (96%) reported a high level of experience using email, 21 (88%) working with word processing applications, 22 (92%) with web browsing, eleven (46%) using drawing applications, and ten (42%) playing games. Fewer had a background using map technologies, where only seven (29%) reported a high level of experience using map software and/or global positioning systems (GPS). Even fewer subjects had experience using small screens. Twenty-one (88%) of all subjects reported that they had little to no experience using a tablet PC. Fifteen (63%) reported little to no experience using handheld computers. Additional information of interest was handedness. Three subjects were left-handed and the other 21 were right-handed.

4.3.3 Experimental task

The last portion of the address canvassing task that involves verification of target addresses and corrections to the map was selected as the basis of our experiment. It is a simple task that involves visualization and orientation components.

The task involved comparing a housing unit configuration on the ground with information in the map. Possible outcomes included: 1) the ground situation was correctly reflected in the map; 2) the map had an error of commission with a map spot that needed to be removed; 3) the map had an error of omission where a map spot needed to be inserted; and 4) the map had an error of wrong position (another error of commission) where a map spot needed to be relocated. After identifying which of the four outcomes applied to a setting, the subject was informed how to take corrective action; i.e., the subject removed a map spot if the outcome was 2), or added a map spot if the outcome was 3). In the instance that a map spot needed to be relocated (case 4) subjects were trained to do the following procedure: 1) add a new map spot and 2) delete the old wrongly positioned map spot.

After a careful analysis of the task, four to five steps were identified that were specific to the task. These steps were 1) find the address on the ground, 2) locate the address vicinity on a map, 3) answer a question as to whether or not the address was on the map, 4) answer a question as to whether or not the address was in the correct location, and 5) fix the map. In instances where the address was neither on the ground nor on the map steps four and five were not required.

After having selected and analyzed the task, we created a design that simulated the environment for the task. The design was determined through iterative testing of storyboards (Appendix F) and flow charts (Appendix G). A software package was developed to instantiate the experiment task. In this software package, the subject viewed a street using photographic images displayed on two monitors, one for each side of the street (Figure 9). In addition to the monitors the subjects were presented with a map-based interface on the tablet PC. Treatments included 1) an interface that provided step-by-step written guidance on what

the user should be doing, and 2) software that had the same functionality, but provided no

guidance (Figure 10).

Figure 9: The computer set up included tablet PC that was situated directly in front of the user and two 20 inch monitors that were 25 inches in front of the user.
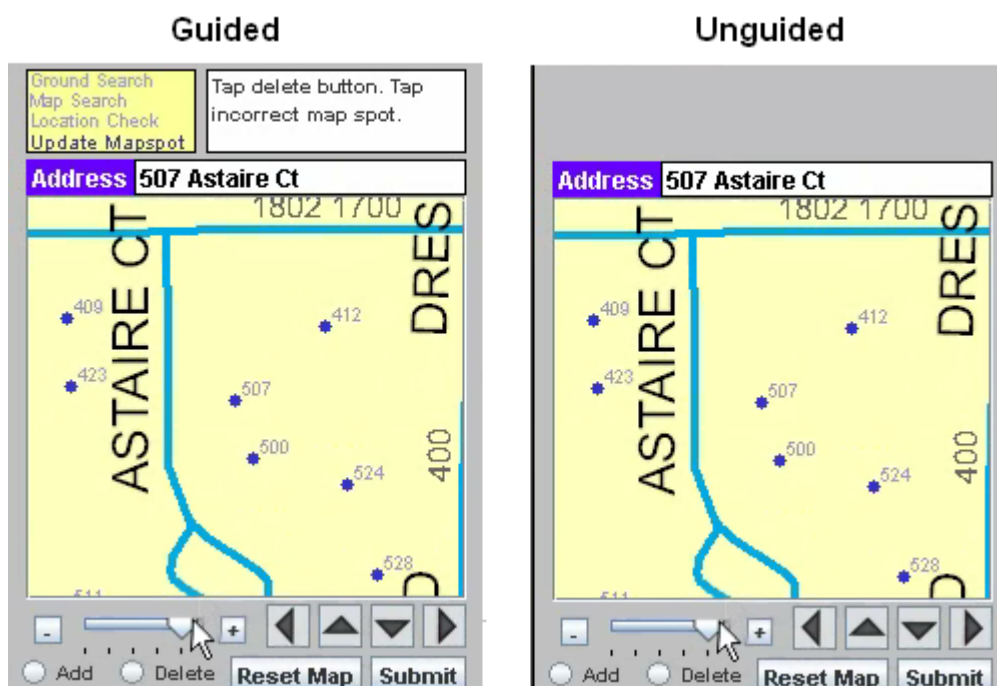
Figure 10: The guided and unguided interfaces.

The guided version included a yellow box at the top of the screen that provided real-time feedback on what stage the user was at. There was a list of stages that the user had to accomplish to complete each scenario. As the user progressed through each stage, the current step that they were on for each screen was highlighted within the list. For example, if the user was at a screen where they needed make changes to the map, the yellow box highlighted the list item "Update Mapspot." Additionally, there was an instruction box on the guided version of the software. The instruction box provided information to the user about what actions needed to be accomplished on each screen. For example, if the user was on a screen in which they were required to fix the map, the screen would tell them one of the specific fixes that needed to be accomplished such as "Tap delete button". Map related functions that were available on both interfaces included zoom, pan, reset map, add map spot,

and delete map spot. The interface also included an address bar that presented a unique target address to the user for 10 different scenarios. The unguided version of the software was identical to the guided version with respect to size and layout, except that the feedback box and instruction box were not included.

The software recorded each action so that a summary of performance measures could be used for analysis. Specific variables of interest included time spent on each screen, attempts in answering address matching questions, accuracy in fixing maps, and tool usage.

4.3.4 Materials

The background questionnaire was administered during the first large group session. This questionnaire included questions related to computer experience, student and employment status, gender, handedness, and the means by which they heard of the study. The questionnaire was identical to the one used for the study described in Chapter 3 with the exception of one additional question regarding handedness (Appendix C). Three cognitive tests were also administered in the large group sessions. These tests included Ekstrom et al. (1976) assessments on visualization and logical reasoning and the Kozhevnikov et al. (2006) perspective taking assessment on orientation. Details on each factor and specified test are as follows:

Visualization—Paper-folding test. The test includes a series of paper-folding problems. The subject sees a folded piece of paper that has one hole punched within it. Their task is to determine where the holes would be located in the paper if it were to be completely unfolded (Appendix D). This test involves 20 problems and is administered in two parts. The full test requires six minutes (three minutes per part).

Logical reasoning—Inference test. This test involves a task where the subject is presented with a statement in which they are instructed to select a valid conclusion from a list of five possible choices. The test is administered in two parts. Each part consists of ten items. The full text requires 12 minutes (six minutes per part).

Orientation—Perspective taking test. The test involves a computerized application where a display shows a representative figure that is situated within an arrangement of landmarks. The subject taking the test is told that s/he should consider her/himself in the position of the representative figure that is in the display. The representative figure is facing in the direction of one of the landmarks. After a few seconds a red flashing spot begins to flash on one of the other landmarks (i.e., the target). The task is to determine the pointing direction of the target landmark. Then the subject must make a mental rotation and determine the pointing direction onto a separate mapping of arrows. A further description of this test is included in Appendix D. The test includes six practice problems and 72 actual test problem and takes approximately 10-15 minutes to complete.

Subjects completed all three cognitive tests in the first session and completed the software exercise in a second session. The two monitors used in the second session to display the street photographs were Dell UltraSharp 2000FP 20-inch Flat Panel Monitors (16 inches in width and 12 inches in height). The subject used the software on a tablet pc that displayed a map of the same area with housing units designated by small symbols (called address or map spots) along the streets. The physical dimensions of the map software on the tablet PC were reduced to emulate the size of a handheld. The specific measurements were 2 ¼ inches in width by 3 inches in height for the active interface area and 2 1/16 inches in

width and 1 7/8 inches in height for the map display area. The computer used to display the interface was a Gateway Tablet PC M1300. This tablet had a 12.1-inch active matrix LCD color display (9 ¾ inches in width and 7 ¼ inches in height). The tablet was configured so that the display was in a landscape setting for the experiment (Figure 9).

For the photographic images, we used manipulated photos of streetscapes (Appendix H). The experiment used maps that were compiled based on Iowa data from Black Hawk County and the Department of Transportation (DOT). These maps were similar to TIGER/Line shape files that are used by the Census Bureau.   The manipulation created settings that challenged the users in ways that were consistent with the objectives of the study. For example, we removed a structure from a photo to create a vacant lot on the ground where the map included an existing map spot.

Six factors were used to differentiate each scenario. These factors included photo, street name, road configuration (e.g., four-way intersection, three-way intersection, etc.), rotation (e.g., north up, south up, etc.), map, and corrective action required (Appendix H). When scored for complexity each category received a score of either one (low complexity) or two (high complexity).

4.3.5 Experiment procedure

In the first session subjects were tested in groups of 2-5 people (with the exception of one person who was scheduled individually because of timing conflicts). After all of the subjects arrived for the first session they were greeted and then presented with an informed consent form (Appendix I). After having read this form and signed it, the paper-based cognitive tests were administered in a conference room. A test script was used to ensure

consistency (Appendix J). The Inference Test on reasoning was administered first and the Paper-folding test was administered second. After the paper-based tests were completed, the subjects were taken to a computer lab where they completed the background questionnaire. Next they were trained on the Perspective Taking Assessment (PTA) and then they proceeded to complete the test. After completing the PTA, test the subjects were excused to leave. Subjects also scheduled a time slot in which they would return for the second session during their first visit. Overall, the second session lasted approximately one hour. The PTA results were compiled by Rusch's adviser Nusser, who created a randomization system for assigning the guided and unguided treatment used in the second session.

The second sessions took place throughout the two weeks that followed the first session. When a subject returned, they were first greeted and then informed about the exercise that they would complete in the second session. The subject was informed that they would perform a task that involved finding a target housing unit on the ground and on the map. Subjects were then trained on an example scenario that was based on two color paper printouts (Appendix J). One color printout included example street photos and the other printout included a zoomed-in map. Next the overhead light was turned out and a smaller lamp was turned on in the corner of the room. The overhead light was turned out to minimize glare on the tablet PC. The next step involved calibrating the tablet touch screen. Calibration was performed by each subject to ensure that the tablet was sensitive to the user's handedness and the way in which s/he used the stylus.

After having completed calibration, the user was trained on two computerized practice scenarios. The user then proceeded to complete 10 test scenarios. After having

completed all 10 test scenarios received a $30 gift card to HyVee or Target according to their preference. After they signed a receipt of payment they were then excused to leave.

4.3.6 Analysis

The impact of interface treatment (i.e., guided or unguided) and associations with demographic and cognitive ability covariates on the subjects' behavioral and performance measures were evaluated using regression. Response variables included time required to perform each scenario, accuracy of locations for addresses that required adding map spots on the map, number of times the pan button was used, and number of times the zoom button was used. Since coordinates for map spots were set as the centroid of the parcel, accuracy of a newly placed housing unit was derived by computing the distance (in meters) between the centroid of the parcel in which the housing unit was located and the location of the housing unit inserted by the subject. Preliminary analyses indicated that the location accuracy variable required a transformation to meet regression analysis assumptions, and thus a log transformation was applied to this variable prior to fitting regressions. The interface treatment variable was expressed as an indicator variable indicating whether the subject was assigned to the guided treatment or not. Demographic variables included in the model were age category (18-29 years of age, 30-39 years of age, 40-59 years of age, or 60 years and older) and a gender indicator for females (female=1, male=0). Additional covariates included functions of the cognitive ability test scores for visualization, perspective taking, and logical reasoning. In particular, for spatial ability, we created two new variables for the regression. First, we used an average of standardized scores for visualization and perspective taking tests as a measure of general spatial ability. In addition, we calculated the difference

between these two standardized scores (visualization – perspective taking) as a measure of the difference between the two types of spatial ability thought to be particularly relevant to our problem. Regression models were fit using PROC GLM in SAS. We examined residuals for departures from assumptions of homogenous variance and linearity. Tests of whether regression parameters were equal to zero were conducted to identify which factors were associated with each response variable.

4.4 Results

Table 10 presents the test results from the analysis of covariance for time, log accuracy, zoom button usage, and pan button usage. There was a significant effect of the standardized average of visualization and perspective taking on time. The slope was -320 (SE=145) which meant that for every unit increase in the average of the visualization and perspective taking standardized test scores, there was a drop of approximately 320 seconds in the average time spent on completing the full exercise.

For accuracy, there was a significant association between error in placing housing unit locations (log meters) and dissociable differences in spatial ability (i.e., the standardized difference between visualization and perspective taking). The slope was -.69 (SE=.26) which meant that for every unit increase in the difference between visualization and perspective taking standardized test scores there was drop of approximately .69 log meters in accuracy. That is, subjects with stronger perspective taking ability had more error in placing map spots relative for subjects of stronger visualization ability who tended to be more accurate when adding new housing unit locations to the map. Additionally, there was a significant

association between gender and accuracy. The slope was 1.25 (SE=.48) which indicated that females tended to be 1.25 log meters less accurate than males.

Table 10. Fixed effects[1] table for time, log accuracy, and tool usage.

Effect	Time (sec)		Accuracy (log m)[2]		Zoom (No. zoom actions)		Pan (No. pan actions)	
	F	p	F	p	F	p	F	p
Interface	0.50		0.15		2.23		0.08	
Age	1.58		0.06		5.79	0.03	1.88	
Gender	0.58		6.92	0.02	3.00		0.46	
Age*interface	0.51		1.00		2.15		0.10	
Gender* interface	0.01		2.32		2.67		0.05	
Avg(Visualization & Perspective taking)	4.89	0.05	0.32		0.88		0.24	
Avg(Visualization & Perspective taking)*interface	0.00		0.32		0.00		2.71	
Visualization – Perspective taking[4]	0.37		7.16	0.02	0.04		6.83	0.02
Visualization – Perspective taking*interface	2.71		0.50		0.13		2.88	
Logical reasoning	0.49		3.53		0.20		1.51	
Logical reasoning*interface	0.04		3.17		0.03		0.35	

1 Numerator degrees of freedom were equal to 1 and denominator degrees of freedom were equal to 12 on all factors.

2 A log transformation was applied to the response variable accuracy.

3 Represents the average standardized score of visualization and perspective taking abilities.

4 Represents the difference between standardized scores of visualization and perspective taking abilities.

Two aspects of tool usage were included in the analysis relating to the use of the zoom and pan buttons. There was a significant association between age and the use of the zoom tool. The slope was -5.18 (SE=2.15) which meant that older subjects tended to make less use of the zoom tool. There was a significant association of the differences between visualization and perspective taking scores and use of the pan buttons. The slope was -52.13 (SE=19.94) which meant that subjects with high visualization relative to perspective taking scores tended to make less use of the pan tool.

4.5 Discussion

An important goal of this research was to investigate the relationship of spatial ability and user performance using pen and map-based applications in relation to two specific sub-factors of spatial visualization and orientation (i.e., perspective taking) abilities. Our results indicated that higher spatial ability scores, as measured by the average of visualization and perspective taking sub-factors, were correlated with faster performance times for tasks that involved both visualization and perspective taking abilities. This association between spatial ability and user performance is consistent with findings from a large body of literature in software use (Dahlbäck, Höök, & Sjölinder, 1996; Egan, 1988; Vicente, Hayes, & Williges, 1987; Egan & Gomez, 1985). In addition, our results extend this finding to map-based interfaces.

We also obtained evidence for differential effects of spatial ability sub-factors corresponding to visualization and perspective taking. Subjects with higher visualization scores relative to their perspective taking scores were better able to accurately record the location of addresses that were missing from the map. Accurately inserting a new map spot relied upon the capacity to maintain a good visual memory while performing a mental rotation. Specifically, the visual memory involved spatial attributes for other map spots and streets in the immediate vicinity. The tests for measuring visualization and perspective taking both required the storage of a visual memory in addition to the ability to mentally rotate objects. They were, however distinct in the way that they placed demands on memory. Specifically, the visualization test required the most serial operations and additionally it placed the largest demand on cognitive processing because it required mental effort to

maintain a visual memory (i.e., the location of the punched holes) in short-term memory while performing multiple serial operations. Thus we considered the steps involved for the visualization test most comparable to the steps involved with adding spots to the map.

Pan usage was also lower for subjects with higher visualization scores relative to their perspective taking scores. Definitive explanations for this relationship are not available from the data we collected. A speculative explanation for this finding is that subjects with low visualization struggled to find the target address on the map when the monitor display was not northbound. In this case their lower spatial skills may have hindered them from recognizing the actual intersection at the closest levels of zoom. As a result they zoomed to the designated vicinity, but because the configuration was not perfectly aligned with the photos they would often defer to panning to see if they should be looking at a different intersection. Another possible explanation is that persons with higher perspective taking scores got confused while attempting navigation within the map because they were trying to rely on their egocentric point of view.

Another finding was that older subjects used the zoom tool less frequently. One reason for this finding may be that older subjects were less comfortable using the zoom tool due to visual impairments that typically occur with aging. Thus because of the tools small size they avoided frequent use of it by simply advancing to the closest level of zoom right away. This explanation while plausible is only speculative. We do not have a lot of insight into the true reason behind the lack of usage thus no conclusions were made.

Spatial ability has also been found to vary by gender, and when this factor is significant, results indicate that spatial ability tends to be higher for men relative to women (Linn & Peterson, 1985). We found that on average, male subjects more accurately placed

housing units on the map when the ground situation showed a housing unit that was not initially present on the map.

Finally, although we were interested in whether guidance on the protocol would mitigate performance issues for subjects of lower spatial ability, we did not find any evidence to support this hypothesis.

CHAPTER 5. CONCLUSIONS

5.1 Introduction

This thesis explored the association between individual differences and software user performance. The primary goal was to investigate the relationship between spatial ability and user performance in the context of map software. Two hypotheses were evaluated in each of our two user studies. First, we expected that performance measures such as time to complete a task or accuracy would be positively correlated with spatial ability, as measured by scores on cognitive assessments for sub-factors of spatial ability. Second, we hypothesized that differences between subjects with low spatial ability and subjects with high spatial ability would be more pronounced with complex tasks and interfaces, such as a map interface or a task to compare ground settings with a map, relative to simpler tasks, such as finding a button that matched a target color or using map software with guidance on how to use the software to execute the task protocol.

Each user study also incorporated additional questions unique to the study. Our first user study investigated how the relationship between spatial ability and performance might be affected by the relevance of the button label and alternative button layouts. The second study investigated whether the association between spatial ability and user performance was affected by offering guidance on performing a task and whether orientation and visualization were dissociable factors for a task involving ground to map comparisons.

5.2 Spatial ability and user performance

The primary goal was to investigate the relationship between spatial ability and user performance in the context of map software interfaces on pen-based field computers. Although prior investigations have consistently demonstrated that software user performance is linked with spatial ability, little research has been conducted in settings that involve small screens and map-based interfaces. Both of our studies demonstrated that user performance was related to spatial ability for the pen-based interfaces under consideration and that for some measures, this relationship depended on the complexity of the task.

5.2.1 Study one: Button labels and layout

For our first study, we investigated relationships between spatial ability and performance in the context of informative and uninformative button labels and alternative layouts of the buttons. We focused in particular on the primary constructs of spatial ability (spatial visualization, perceptual speed) and related components of spatial visualization (visual memory).

Spatial visualization scores were negatively associated with the time it took subjects to complete the initial block (i.e., initial time) and the accuracy of button location recall regardless of interface type or button label, which indicated that the inexperienced performance appeared to be mediated to some extent by spatial visualization ability. This was a setting that might be considered challenging because users would have been unfamiliar with the task at the beginning of an exercise.

In some cases, the relationship between spatial visualization and performance was only present for complex layouts. The association between total time to complete the

109

exercise and spatial visualization scores held for the map-based interface, but not for the simpler color-matching tasks. Similarly, total exercise time was negative associated with perceptual speed scores for grid and map-based interfaces, but not with the circular layout. A negative relationship was found between visual memory and total attempts required to complete an exercise for the arbitrary label. This relationship was stronger for the circle interface than it was for the grid and map interfaces, perhaps due to the lack of cues offered by the radial symmetry of a circular layout. These results all substantiate our hypothesis that spatial ability is likely to have a stronger association when interface layouts are more complex.

Interactions between spatial ability and button label also indicated the importance of spatial ability when the task or setting was difficult. The negative relationship between visual memory and total time was only significant when labels were uninformative, where subjects were provided with minimal assistance in associating button functions with labels. A negative relationship between total attempts and visual memory was evident only for the arbitrary graphic, which apparently presented a challenge to the user that exceeded the difficulty of remembering the location of unlabeled buttons.

This work points to the importance of using appropriate labels in software design. Our results suggest that differences in performance due to variation in spatial ability may be mitigated to some extent by having informative labels. Choosing appropriate icons is simpler for software functions that are commonly used because familiar icons or conventions are already available (e.g., editing functions in word processing software, icons for saving a file). In contrast, when creating applications that involve unfamiliar tasks, such as complex map functions, icon selection can be more challenging. It is important to build on previous

associations with the task to the extent possible; when this is not possible, user testing should be used to identify labels that users will more readily recall.

5.2.2 Study two: Software guidance and spatial visualization versus perspective taking

The second study investigated the impact of spatial ability in the context of guided and unguided map-based interfaces. The task and the software interface were both considered complex, and offered a more realistic setting in which map software would be critical to performing a task.

The study confirmed results found in the first study that subjects with lower spatial ability (as measured by the average of visualization and perspective taking) tended to require more time to perform a task when using map-based software than subjects with high spatial ability. There was no relationship between accuracy in performing a location-based task and the average of the two spatial ability measures, although as discussed below, there appeared to be an indirect indication that a relationship existed for components of spatial ability.

One area of interest was whether spatial visualization and perspective taking would have a differential association with specific performance measures. We found that subjects whose perspective taking scores were higher than their spatial visualization scores tended to provide less accurate representations of housing unit locations on the map and make more use of the panning tool relative to persons whose visualization scores were higher than their perspective taking scores. We hypothesized that having better perspective taking relative to visualization abilities allowed subjects to do a better job of aligning the photographic ground and map representations. However, their lower visualization abilities may have resulted in

the need for more map display manipulations to identify the correct location on the map than subjects whose visualization scores exceeded their perspective taking scores.

The effectiveness of software guidance remains an open question. Although we had hypothesized that guidance would help mitigate differences between subjects with low and high spatial ability scores, we did not find evidence to support this conjecture. The lack of any interface effect in our second study was likely due to the ineffectiveness of the guidance mechanism in the context of a map-based application. Much of the literature on guided applications proposes that when cues are integrated within the content that is being processed, the cues are most likely to be effective (e.g., Kalyuga et al., 1998).

5.3 Implications for Census Bureau applications

The Census Bureau is considering the use of map software interfaces for field activities such as address canvassing. Our research was designed in part to explore fundamental interface design issues for this setting. We focused on software that offers digital maps on a handheld computer to support field verification and updating of the Bureau's address lists, which are developed for the decennial census and for constructing lists to select samples from. Although tools such as Google Maps have led to increased familiarity with manipulating geographic information for a broad range of our population, open questions remain about how efficiently people use these tools and how accurate their work is. In particular, because prior research had identified the importance of spatial ability in effectively using interfaces and in how people mentally work with geospatial information, our goal was to explore the relationship between spatial ability and performance when using map-based interfaces.

112

Our research suggested that variation in spatial ability will affect at least the efficiency with which Census takers perform location-based tasks in this kind of setting. Efficiency measures included time to perform a task and patterns of button use, such as number of attempts required to complete a task correctly or number of uses of specific map function tools. Software efficiency is an important consideration in keeping costs down for field surveys, particularly in the decennial census setting where hundreds of thousands of field enumerators use the software.

We did not find direct evidence to support a relationship between spatial ability and the accuracy of new location. This may have been due to the small sample size in our second study, which used a task that was closely related to field address verification. In our first study, subjects recalled and marked the locations of software button positions on an interface. Results from this study, which involved subjects with a greater range of spatial ability scores and a larger sample size, suggested that the accuracy of a location task tended to be poorer for persons with lower spatial ability scores.

This research underscores the importance of interface design in reducing costs and burden for users of map-based software in a small screen computing environment. One way interface design can be used to minimize user burden and costs is to develop tools that accommodate for users of a broad age range of spatial abilities. Further research is needed on exactly how this might be accomplished. However, our research indicated that good button labeling helps to mitigate performance differences among subjects with varying spatial ability. In addition, it seems that tool use strategies may vary with spatial ability. More research is needed to identify patterns of tool use in manipulating maps (e.g., ways in which pan and zoom are used) and to understand why users exhibit particular patterns.

113

Another possible consideration is to use spatial ability tests as a pre-screening device to assess training needs, although this may be more applicable to non-governmental survey organizations. The results from this study showed that the spatial visualization test was reasonably sensitive to variation in the effectiveness of users in performing location-based tasks. The Census Bureau or other survey organizations that use map-based software in the field might develop training content that specifically targets individuals who have more difficulty using maps (regardless of whether a screening test is used).

5.4 Limitations and future work

This research was designed to provide insight into the role of spatial ability in user performance when working with small screen pen interfaces that might be used in location-based field operations, with a special focus on map-based interfaces. Our first study investigated whether the positive association between effective performance and spatial ability that has been found in desktop software settings also applied in a pen-based setting with a smaller interface that involved the complication of map content. While this study confirmed these findings and highlighted the importance of button labeling for map-based tools that are less familiar to users, the functionality of the map software was limited to a static map display in response to a button selection. A more realistic setting would have involved a sequence of map tool actions to execute a specific task. Because we found that spatial ability tended to be more important in complex settings, we would expect results in more realistic settings to confirm findings from our first study. The software and task setting in our second study allowed us to explore these relationships based on an actual address canvassing task with functional map software. However, the experiment was conducted in a

laboratory setting with photographic displays of the environment rather than outdoors with the ability to navigate and gain a fuller perspective of the surrounding environment. An outdoor setting would be more reflective of the actual field environment, but conducting field studies is far more expensive and places a heavier burden on subjects. Again, because of the complexity of an outdoor environment relative to our laboratory setting, we would expect our findings to be supported in a field setting.

The second study also involved a small sample size due to resource constraints. This not only limited the power to detect differences among treatments, but limited the range of spatial ability scores observed. If we had been able to recruit and test a larger subject pool, we might expect results to be more nuanced (as they were in the first study).

An additional difficulty was the lack of true integration of the guidance into the software functions due to resource constraints in developing the study software. The role of guidance in mitigating performance differences remains an open question, and it would be useful to conduct an additional study to evaluate whether guidance would help persons with lower spatial ability when more effectively integrated into the software interface,

Results reported in this thesis are limited to coarse summaries for each user, but data exist on performance for specific photographic scenarios and on sequences of tool use. A future research area is to conduct a more detailed examination of patterns of tool use in relation to the complexity factors for each scenario level and spatial ability scores of subjects. Self-reported strategy and user satisfaction data have also not been fully explored.

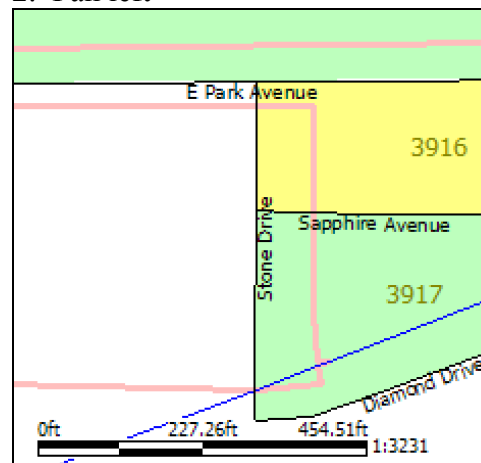# APPENDIX A. EXPERIMENT ONE - MAP VIEWS
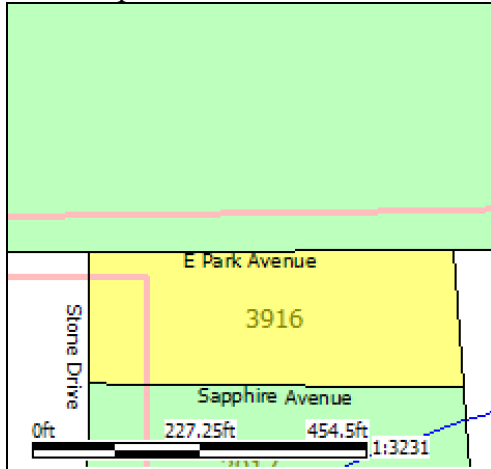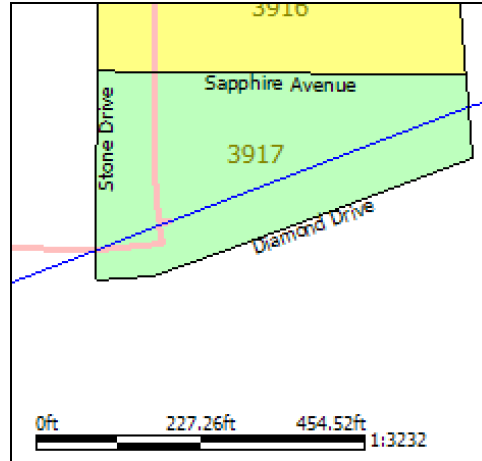
Map views generated by each button function.
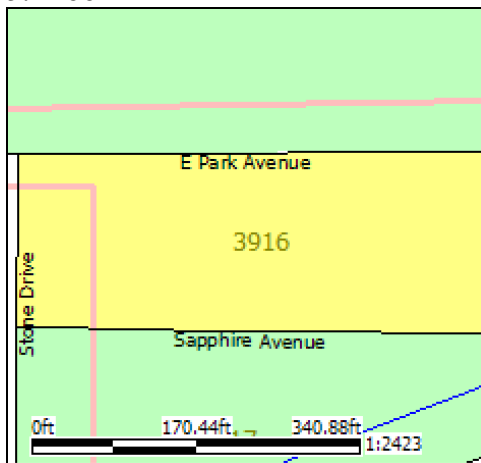
Base view
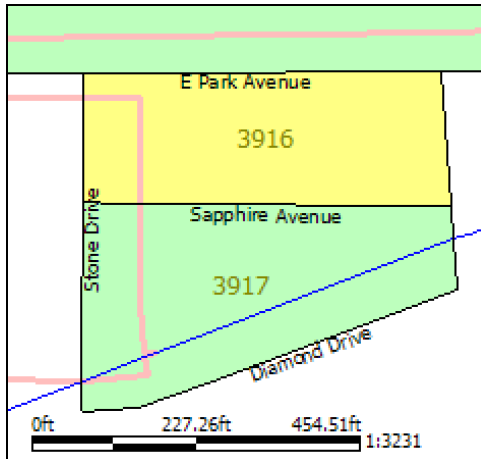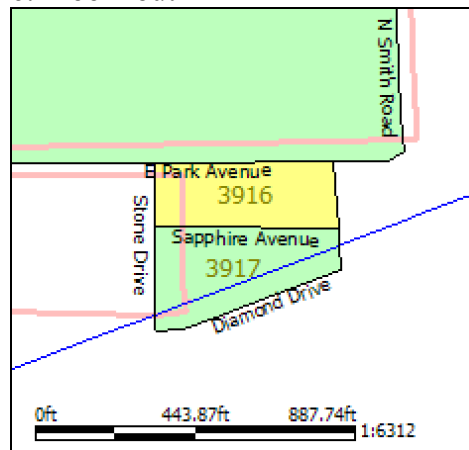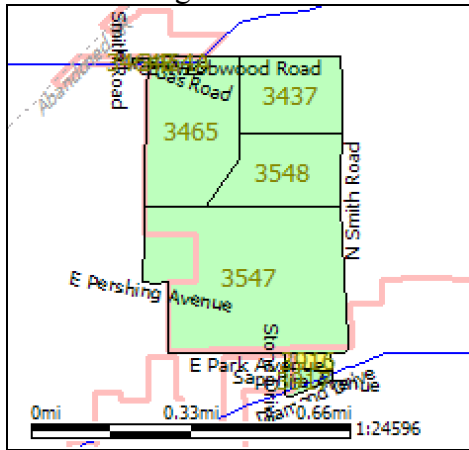

1. Pan right
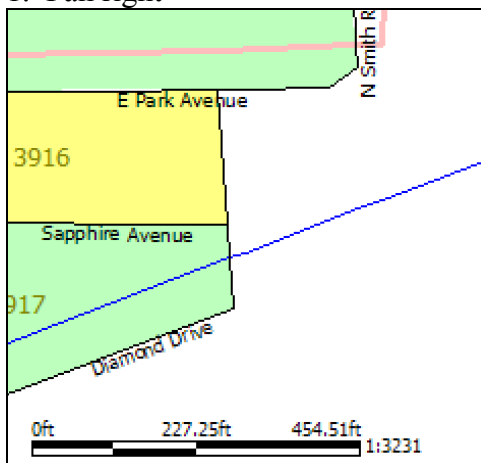
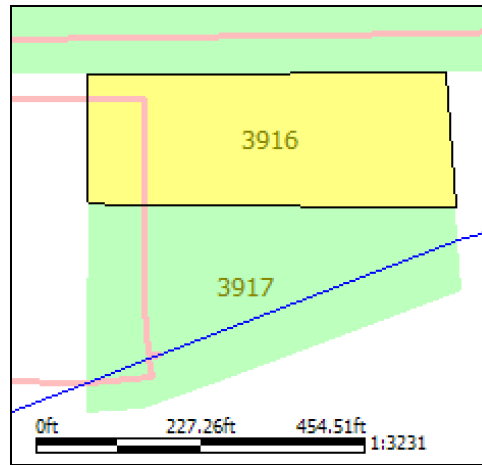2. Pan left

3. Pan up

4. Pan down
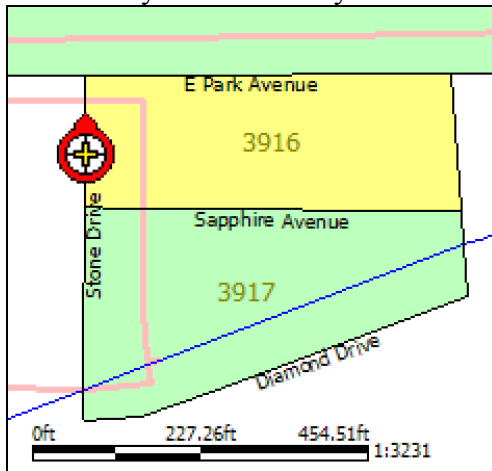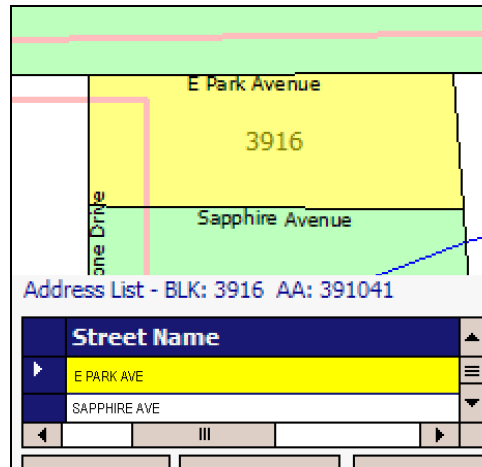

5. Zoom in

6. Zoom out
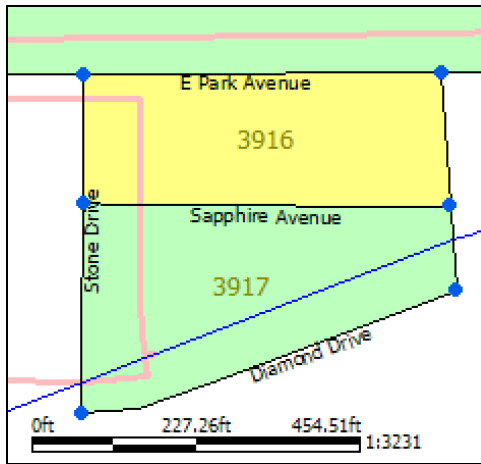
7. Show assignment area

8. Remove street names

9. Show "you are here" symbol

10. Show street list

Address List - BLK: 3916 AA: 391041
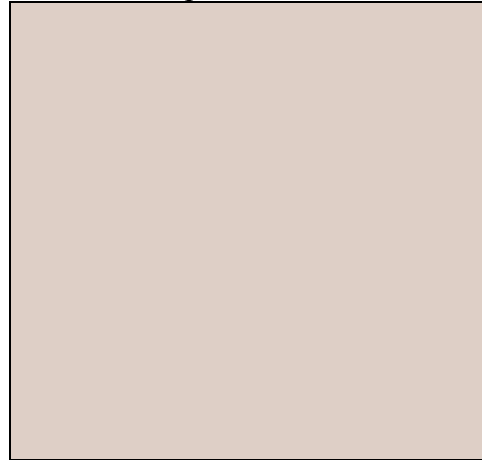
Street Name
E PARK AVE
SAPPHIRE AVE

11. Show intersections

12. Close map

APPENDIX B. EXPERIMENT ONE – INFORMED CONSENT

Letter of introduction with elements of consent

Software interface design

Our goal is to study software interface design on tablet PCs.

The exercise will take about one hour to complete. During the study you may expect the following procedures to be followed. You will complete a short questionnaire, a few color training and cognition assessments, and three exercises using a software interface.

There are no known risks to participation other than ergonomic concerns that are normally associated in the usage of the tablet PC.

There are no direct benefits to you as a participant other than the educational experience of being involved in an experiment. Your participation is helping the United States Census Bureau to develop better software.

If you complete this study in full, you will be provided with a $25 gift card for your participation.

Your participation in this study is completely voluntary and you may withdraw from the study at any time.

Records identifying participants will be kept confidential. Results will be released in summary form only.

Only researchers from Iowa State University working on this project will have access to data collected during this study.

The data collected in this research may be used for educational or scientific purposes and may be presented at scientific and/or educational meetings or published in professional journals. Published results will be in summary form only with no personal identifying information.

For further information about the study contact Michelle Rusch or Dr. Sarah Nusser at (515) 294-9773. If you have any questions about the rights of research subjects or research-related injury, please contact Diane Ament, Director, Office of Research Assurances (515) 294-3115, dament@iastate.edu.

APPENDIX C. EXPERIMENT ONE – BACKGROUND QUESTIONNAIRE

Background information

Please answer each of the questions by writing your answer in the space provided, or by circling the number next to your selected answer.

1. How much experience do you have with the following computer and software tools?

	No experience		Moderately Experienced		Very Experienced
a. Personal or laptop computer	1	2	3	4	5
b. Tablet PC	1	2	3	4	5
c. Handheld computer	1	2	3	4	5
d. Map software (on your computer or Internet)	1	2	3	4	5
e. GPS (global positioning system) receiver	1	2	3	4	5
f. Word processing applications	1	2	3	4	5
g. Drawing/graphics applications	1	2	3	4	5
h. Web browsing	1	2	3	4	5
i. E-mail	1	2	3	4	5
j. Games	1	2	3	4	5

Please circle the number for all that apply.

2. Student status

    1= Undergraduate student
    2= Graduate student
    3= Not a student

3. Employment status

    1= Employed full time
    2= Employed part time
    3= Retired
    4= Not employed or retired
    Other:_____

4. Age

    1= 18-24 years
    2= 25-35 years
    3= 36-50 years
    4= 51-65 years
    5= 66 years or older

5. Gender

    1= Male
    2= Female

6. How did you hear about this study?

    1= Newspaper
    2= Email announcement
    3= On campus flier at Iowa State
    4= Temporary employment services office
    5= Word of mouth
    Other:_____

APPENDIX D. EXPERIMENT ONE AND TWO – COGNITIVE TEST

EXAMPLES EXPLAINED

Spatial Visualization (Ekstrom, French, Harman, & Dermen, 1976)
Paper Folding Test – VZ-2
"Suggested by Thurston's <u>Punched Holes</u>. For each item successive drawings illustrate two or three folds made in a square sheet of paper. The final drawing of the folded paper shows where a hole is punched in it. The subject selects one of 5 drawings to show how the punched sheet would appear when fully reopened. Length of each part: 10 items, 3 minutes; Suitable for grades 9-16" (p. 174).

Perceptual Speed (Ekstrom et al., 1976)
Number Comparison Test – P-2
"Suggested by the <u>Minnesota Vocational Test for Clerical Workers.</u> The subject inspects pairs of multi-digit numbers and indicates whether the two numbers in each pair are the same or different. Length of each part: 48 items, 1 ½ minutes; Suitable for grades 6-16" (p. 124).

Visual Memory (Ekstrom et al., 1976)
Building Memory – MV-2
"The subject is asked to indicate the location of a number of buildings seen on a previously studied map. Length of each part: 12 items, 4 minutes for memorizing, 4 minutes for testing; Suitable for grades 6-16" (p. 110).

Logical Reasoning (Ekstrom et al., 1976)
Inference Test – RL-2
"A test suggested by a similarly named test by Guilford. The task is to select one of 5 conclusions that can be drawn from each given statement. Length of each part: 10 items, 6 minutes; Suitable for grades 11-16" (p. 142).

Orientation – Perspective taking (Kozhevnikov, Motes, Rasch, & Blajenkova, 2006)
" In the perspective-taking version of the pointing direction task, a small figure representing a character's head indicated the starting location, where participants were to imagine themselves standing […]. The character's eyes were looking towards one of the five locations that represented the to-be-imagined facing location (imagined orientation). Participants were to indicate the direction to a third (target) location from the imagined orientation by pressing a corresponding key on the computer keyboard. Instructions appeared at the bottom of the screen, for example […], 'Imagine you are the figure. You are facing the University. Point to the Airport'. Thus, participants were to imagine transforming their actual perspective (i.e. an aerial perspective of the character and the town) to that of the character's perspective, and then the participants were to imagine pointing to the target from the character's perspective" (p. 400).

APPENDIX E. EXPERIMENT ONE – TEST SCRIPT

Protocol test script

Iowa State University

July 6, 2006

1. Recruit participants and make an appointment for them to be involved with the study

2. Randomly assign the four treatments and two exercise orderings independently to the participant

3. Introduce experiment and human subjects information

   Prior to participant arrival, make sure to have the following:
   - Two copies of the letter of introduction with elements of consent
   - A copy of the experimental protocol
   - Three spatial cognition tests
   - Two tablet PCs (Where Tablet 1 includes the color training task, the background questionnaire, and the practice blocks for exercises 1, 2, and 3.)
   - Compensation

   When the participant arrives, welcome him or her into the experiment room.

   Researcher: The purpose of this study is to investigate principles used to design point-and-tap style software interfaces. Before we begin we need to provide you with more specifics about your rights as a participant, which are outlined in this document. Please read it carefully and then let me know if you have any questions.

   Provide letter of informed consent and obtain consent

4. Have both tablets set up to perform a calibration. Determine if the subject is left-handed or right-handed and adjust the tablets accordingly.

5. Conduct color training test. Hand the subject Tablet 1.

   Researcher: In this training exercise, you will be presented a colored rectangle in the center of the application screen. You need to select the name of that color from the pop-up menu to the right.

You will make your selections using the stylus that is attached to the tablet. This can be done by gently tapping the screen, keep in mind that the tablet usually responds better to very light taps. Once you have selected the name of the color tap the OK button below the pop-up menu.

If you have correctly selected the right button you will advance to the next color. Tap the "ok" button to move on to the next color. When another color is presented follow the same procedure. That is, select the name of the color, and tap ok again.

If you should choose an incorrect color, a screen will display that reads "Incorrect, the correct answer is blue (for example)." Tap the "ok" button to move on to the next color.

You will be selecting from 12 colors. Each color will be shown 3 times each for a total of 36 trials. After all the colors have been displayed, a window will appear indicating that you have completed the full test. When this window appears please hand the tablet back to myself. Do you have any questions?

Go ahead and complete the exercise.

Once the participant has completed the exercise retrieve the tablet back from them. Check the log file to determine the success of the test.

Researcher: You have successfully completed the color training test.

6. Administer background questionnaire

Researcher: Now we have a short questionnaire for you to complete.

Researcher enters id number on Tablet 1 at a configuration screen for the background questionnaire and the software is initiated. Hand Tablet 1 back to the participant.

Researcher: There are three screens to this questionnaire. After you have finished each screen, tap on the "next" button until you get the final screen. After you have completed the final screen there will be a button at the bottom of the screen that says "finish." Tap on this "finish" button once you are finished and then we will move on to the next part of the experiment.

7. Administer cognition tests

Researcher: Next, we would like to give a few standardized tests.

Read instructions specific to each spatial cognition test.

   a. Paper folding test (Part 1 and 2; 7 minutes)
   b. Number comparison test (Part 1 and 2; 5 minutes)

c. Building memory test (Part 1; 11 minutes)

Researcher: You will now be asked to perform a series of tasks on the computer. This will involve the use of three different interfaces. First we will step through a few practice exercises.

8. Researcher enters id number on Tablet 2 at the configuration screen and the software is initiated.

Look up the participant assignment list to obtain the condition under which the subject is to be run. This will be pre-determined through the random assignment that is described in step two.

When the configuration screen appears, the researcher will type the assigned number for the particular subject. After the number has been typed they will then press the enter button. When the dialog window disappears and the application window appears everything is ready to go.

9. Participant introduced to first exercise and practices using the specified button interface

Researcher: Go ahead and tap the rectangular button in the center of the window on the screen and we will begin with the practice trials.

You see that after you tap on the centered rectangle twelve buttons appear. Your goal is to find the button that corresponds to the target color [in the center below the two rows of buttons or in the center of the circle of buttons].

One way to determine the color associated with the button is to hold the cursor over the button and activate the tooltip. To do this, it works best to hold the stylus over the button and keep it very still.   Go ahead and move the cursor over a button now. As you can see a little square appears with the button's associated color. If it is not the right one you can just move the cursor over each button until you find the correct one.

Go ahead and look for the button that matches the color in the rectangle. After you have identified the button tap on it. In the case that you select a button that has a different color than the one in the target rectangle a dialog box will pop up that indicates a wrong answer. Go ahead and tap the ok button and then try to select the matching color.

Notice that when you tap on the correct button, all of the buttons disappear. To proceed to the next target color, you will have to tap on the rectangle again. Go ahead and tap the rectangle again.

Go ahead and continue with this task until a message screen appears telling you that you are done. You will go through 12 colors. Keep in mind that this is practice.

The subject will do this task a total of 12 times. Provide guidance if the subject is having problems understanding what to do.

10. Participant executes the first exercise and finalizes it by completing a location test

Researcher: You are now ready for the first exercise. This exercise consists of 16 blocks. After each block, a message will appear on the screen. When these messages appear tap on the ok button and proceed to the next block. When you have completed all 16 block you will see a message telling you that you have finished. At this point in time stop what you are doing and let me know that you are finished.

The main task will be identical to the practice task. The basic procedure for doing the task is the same as the one you just practiced and the button locations are the same. Do you have any questions at this time?

Be sure to work as quickly and as accurately as possible. If you are ready, go ahead and tap the rectangle in the center of the screen to get started.

11. Introduce location test

Researcher: This next exercise involves a test that evaluates your memory for where each button is located. We will now do a short practice exercise so that you will be familiar with how to complete the test.

Tap the "ok" button on the dialog box to go to the practice location test and hand the tablet pc to the participant.

You see on the screen the colored rectangle that was present in the former task. In addition to this rectangle you see a button with no label centered above the colored rectangle. Tap on the button once and you will notice that the button changes color to a dark teal. Move your stylus to drag the button to the location where it was located in the former exercise. Once you have found the location you would like to place the button tap on the screen in that position. When the button changes back to the original gray color you know that it is no longer in the positioning state. If you are not happy with the position you can tap on it again and reposition it. Once you are happy with the location, press the colored rectangle in the center of the screen to advance to the next button.

This was practice, now you will complete a full location test that includes all 12 colors. That is, you will be asked to place the button for all 12 colors. You may go ahead and begin. When you complete the exercise let me know and we will move on to the next thing.

12. Participant takes a short break between exercises

Researcher: You have successfully completed the first exercise. At this time you are free to take a short break before we begin with the next exercise.

13. Participant introduced to the second exercise on Tablet 1 and practices using the specified button interface

    Researcher: This second exercise is similar to the first exercise except for the fact that it uses a different arrangement of buttons on the screen. Go ahead and tap the rectangular button in the center of the window on the screen and we will begin with the practice trials.

    Go ahead and continue with this task until a message screen appears telling you that you are done with the practice blocks.

14. Participant executes the second exercise on Tablet 2 and finalizes it by completing a location test

    Researcher: You are now ready for the second exercise. Similar to the first exercise this exercise consists of 16 blocks. After each block, a message will appear on the screen. When these messages appear tap on the ok button and proceed to the next block. When you have completed all 16 blocks you will see a message telling you that you have finished the exercise.

    The main task will be the same as the practice task. Do you have any questions at this time?

    Be sure to work as quickly and as accurately as possible. If you are ready, go ahead and tap the rectangle in the center of the screen to get started. When you complete the exercise and get a dialog that says "End of exercise" let me know and we will move on to another location test.

15. Introduce second exercise location test

    Researcher: This next practice involves a location learning test again. This test is similar to the location test that you completed for the first exercise. We will now do another short practice exercise so that you will be familiar with how to complete the test.

    Tap the "ok" button to start the location test and hand the tablet pc to the participant.

16. Participant introduced to the third exercise on Tablet 1 and practices using the specified button interface

    Researcher: You will now see a map in addition to a set of buttons. The buttons represent functions that change what you see on the map.

Show them the map screens document and go through each function separately.

Tap on the large white centered rectangle at the bottom of the screen. In response to your tapping, you see that two things have been added to the screen. This includes two rows of buttons that have appeared above the white rectangle and a large box that has a map within it at the top of the screen.

Your task is to read the instruction within the white box and identify the button that matches the instruction. Go ahead and look for the button that you would expect matches the instruction, after you have identified the button tap on it.

One way to determine the function associated with the button is to hold the cursor over the button and activate the tooltip. Go ahead and move the cursor over a button now. As you can see a little square appears that provides a text description of the button's target function. If it is not the right one you can just move the cursor over each button until you find the correct one.

Notice that when you tap on the correct button, the map is updated to reflect the instruction. This change will only be shown for a few moments, after that the map box and the two rows of buttons will disappear. In addition, the white rectangle will once again be blank. To proceed to the next instruction, tap on the white rectangle.

Continue with this task until a message screen appears telling you that you are done. Keep in mind that this is practice.

The subject will do this task a total of 12 times. Provide guidance if the subject is having problems understanding what to do.

17. Participant executes the third exercise on Tablet 2 and finalizes it by completing a location test

Researcher: You are now ready for the third exercise. Similar to the first two exercises this exercise consists of 16 blocks. After each block, a message will appear on the screen. When these messages appear tap on the ok button and proceed to the next block. When you have completed all 16 blocks you will see a message telling you that you have finished the exercise.

The main task will be the same as the practice task. Do you have any questions at this time?

Be sure to work as quickly and as accurately as possible. If you are ready, go ahead and tap the white rectangle centered at the bottom of the screen to get started. When you complete the exercise and get a dialog that says "End of exercise" let me know and we will move on to the final location test.

18. Introduce third exercise location test

Researcher: You see on the screen two familiar objects from the last exercise. This includes the map and the text instruction. In addition to these things you see a button with no label centered above the text instruction. Tap on the button once and then move your stylus to drag the button to the location where it was located in the former exercise. Once you are done placing the button tap the instruction box to advance.

This was practice, now you will complete a full location test that includes all 12 target functions. You may go ahead and begin. This test will be the conclusion of the experiment, when you complete the exercise let me know.

19. Participant given incentive

Researcher: Thank you for participating.

After this, hand them their compensation and send them on their way.

20. Participant excused to go

APPENDIX F. EXPERIMENT TWO – STORYBOARDS

Screen Two

Fall 2007 study
Storyboards
November 8, 2007

Left monitor

Right monitor

Guided

Not guided

Start screen

map_id

Start screen

End screen

End screen

Screen Two A

Left monitor

Right monitor

Guided

Not guided

map_caption_timeout= __ s

map_caption

1

2

Ground search
Map search
Location check
Update mapspot

Hint: Try to zoom out.

Address

3910   Astaire Ave.

Would you like to return to the starting zoom level?

Yes   No

timeout1= __ s

timeout1_prompt

Start screen

map_id

Address

3910   Astaire Ave.

Start screen

Add   Delete   Help   Reset map   SUBMIT

Add   Delete   Help   SUBMIT

Ground search
Map search
Location check
Update mapspot

Address

3910   Astaire Ave.

You have exceeded the time limit. You will return to the starting zoom level.

timeout2= __ s

timeout2_prompt

End screen

Address

3910   Astaire Ave.

End screen

Add   Delete   Help   Reset map   SUBMIT

Add   Delete   Help   SUBMIT

Screen Three

Left monitor

Right monitor

1 Guided

2 Not guided

134

Fall 2007 study
Storyboards
October 16, 2007

Screen Three A

Left monitor
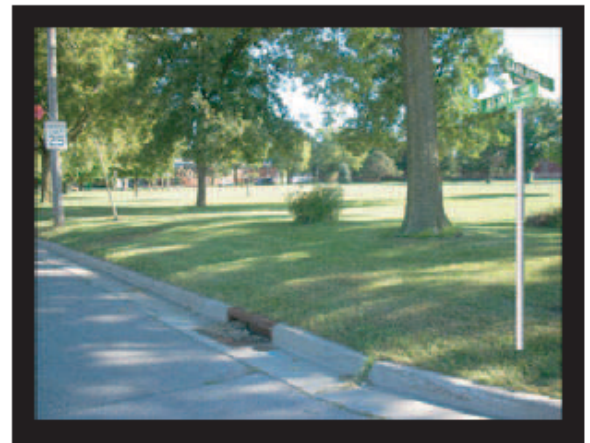
Right monitor
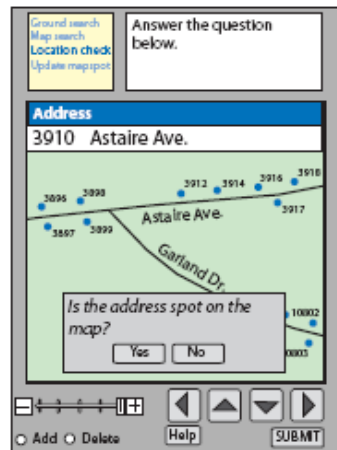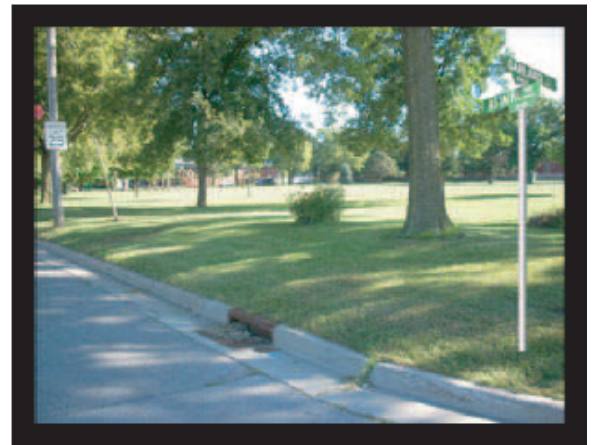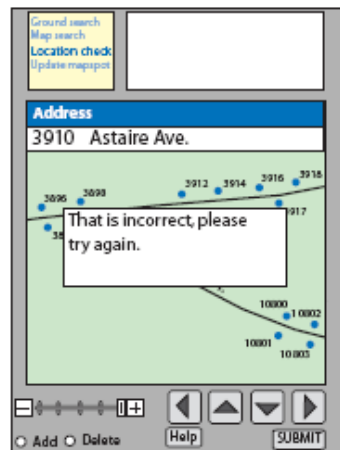
1 Guided
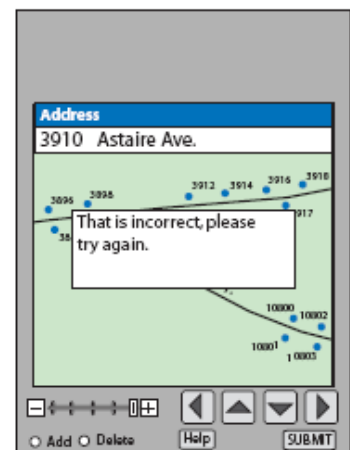
2 Not guided

Screen Four

Left monitor

Right monitor

1 **Guided**

2 **Not guided**

136

Fall 2007 study
Storyboards
October 16, 2007
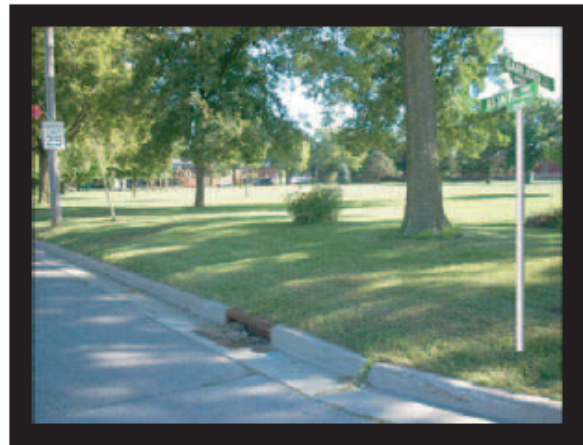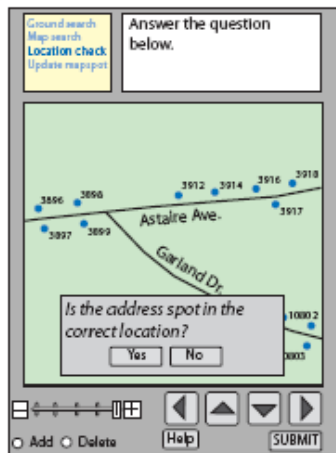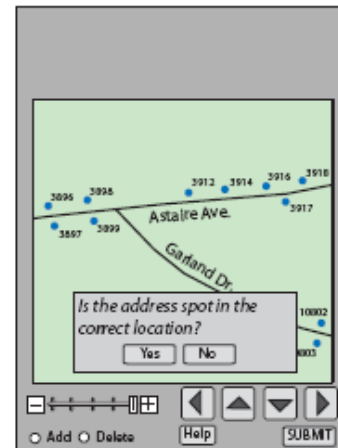
Screen Four A

Left monitor

Right monitor

1 Guided

That is incorrect, please try again.

2 Not guided

That is incorrect, please try again.

www.manaraa.com
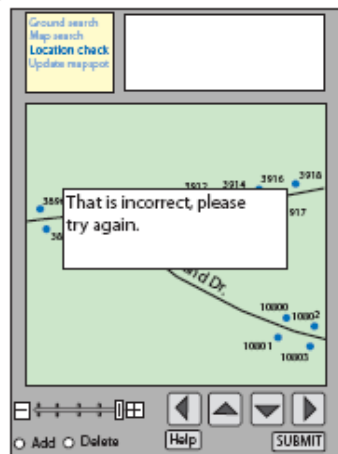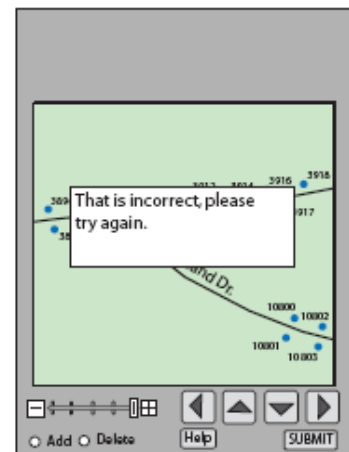
Screen Five

Left monitor
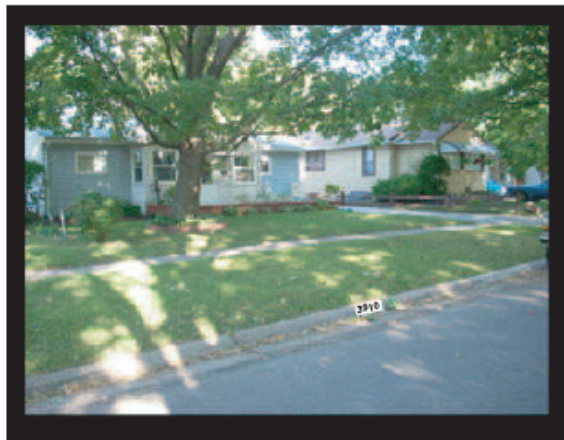
Right monitor
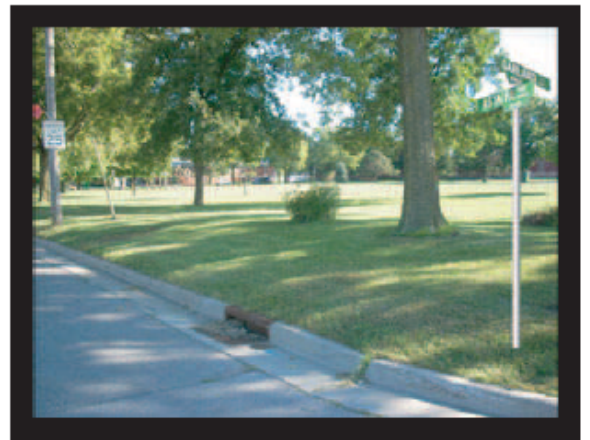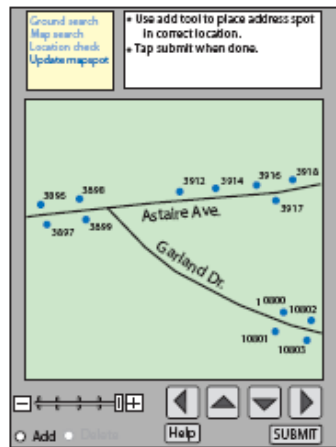
1 **Guided**

2 **Not guided**
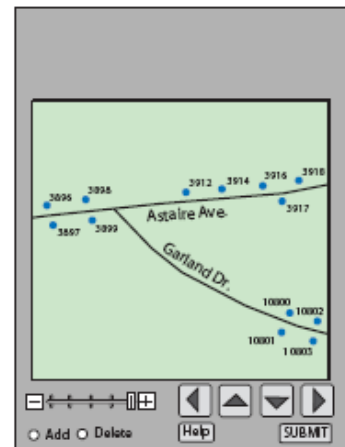
APPENDIX G. EXPERIMENT TWO – FLOWCHART

Experiment two: Task map flowchart 9/26/2008

Screens

Scene 0
Introduction
(Screen 0)

Start

Scene 1
Ground Search:
Address + Ground + Address list
(Screen 1)

Incorrect

Try again
(Screen 1A)

Incorrect

Correct

Correct

Scene 2
Map Search:
Address + Ground + Zoomed Out Map
(Screen 2)

Submit

Yes

Location check (1):
Address + Ground + Zoomed In Map
(Screen 3)

Incorrect

Try again
(Screen 3A)

Incorrect

No

Submit

Yes

Scene 3
Location check (2):
Address + Ground + Zoomed In Map
(Screen 4)

Yes   Incorrect

Try again
(Screen 4A)

3 sec delay

No

Scene 4
Update mapspot:
Address + Ground + Zoomed In Map
(Screen 5)

Submit

Scene 5
Done

Actions

1) User taps "Start" to begin exercise

1) User matches ground to address in list
2) User answers question "Is there a matching housing unit on the ground?" by tapping "Yes" or "No"

Repeat of actions from Screen 1

1) User uses arrows to pan and slider to zoom
2) User might use "Help" option
3) User taps "Submit" when done

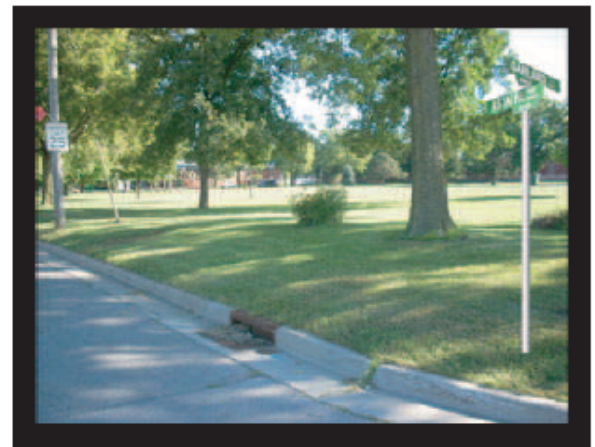1) User answers question "Is the address spot on the map?" by tapping "Yes" or "No"
2)

1) User answers question "do you want to zoom again?"

1) User answers question "Is the address spot in the correct location?" by tapping "Yes" or "No"

1) User taps "Add" radio to place a new spot.
2) User taps "Delete" radio to remove incorrect spot

Data capture

Tap starts *Ground Search* time

Tap of the correct answer ends *Ground Search* time and starts *Map Search* time

Tap "Submit" ends *Map Search* time and starts *Location Check* time

Tap "Yes" ends *Location check* time and starts *Map Search* time again.

Tap "No" starts *Update mapspot* time

Tap "Yes" ends *Location Check* time and trial is over.

Tap "Submit" ends *Update mapspot* time and trial is over.

APPENDIX H. EXPERIMENT TWO – COMPLEXITY TABLE AND SCENARIO PHOTOS

Order	Street1	Street2	Street3	Name complexity	Road type*	Road complex	Photo complexity	Ground orientation
0A	Trent St	1st St		1	Two-way intersection	2	2	1 Northbound
0B	Main St	22nd St		1	Three-way intersection	2	2	2 Westbound
1	Gonda St			1	Dead end	1	1	1 Eastbound
2	F St NE	8th St.		2	Four-way intersection	2	2	1 Southbound
3	Wrightwood St			1	Dead end	1	1	1 Northbound
4	Cardinal Court			1	Culdesac	1	1	2 Eastbound
5	Jones Dr			1	Two-way intersection	2	2	2 Southbound
6	Barnett Rd	12th St	Dressell Rd	2	Four-way intersection	2	2	1 Westbound
7	Ford St	23rd St		1	Two-way intersection	2	2	2 Eastbound
8	E St NE	20th St.		2	Three-way intersection	2	2	2 Northbound
9	Astaire Ct			1	Culdesac	1	1	1 Westbound
10	Promise St	12th St		1	Three-way intersection	2	2	1 Southbound

Photo complexity
1- includes 1-3 address captions
2- includes 4 or more address captions

Map fix codes
1 - Ground: yes
    Map: yes
        Correctly located: yes
        Fix: none
2 - Ground: yes
    Map: yes
        Correctly located: no
        Fix: add new mapspot and delete old mapspot
3 - Ground: yes
    Map: no
        Correctly located: no
        Fix: add new mapspot
4 - Ground: no
    Map: yes
        Correctly located: no
        Fix: delete mapspot
5 - Ground: no
    Map: no
        Correctly located: no
        Fix: none

* There is not an equal distribution of road type (4 types) and rotation (4 types)

Order	Street1	Street2	Street3	Rotation*	Rot complexity	Ground complexity score	Map complexity	Fix complexity	Map fix	Overall complexity score	Easy/hard
0A	Trent St	1st St		0	1	5	1	2	4	8	Easy
0B	Main St	22nd St		270	2	7	2	2	2	11	Hard
1	Gonda St			90	2	5	2	1	4	8	Easy
2	F St NE	8th St.		180	1	6	2	2	5	9	Hard
3	Wrightwood St			0	1	4	1	2	2	7	Easy
4	Cardinal Court			90	2	6	2	1	3	9	Hard
5	Jones Dr			180	1	6	1	1	1	8	Easy
6	Barnett Rd	12th St	Dressell Rd	270	2	7	1	2	4	10	Hard
7	Ford St	23rd St		90	2	7	2	1	3	10	Hard
8	E St NE	20th St.		0	1	7	1	2	5	10	Hard
9	Astaire Ct			270	2	5	1	1	2	7	Easy
10	Promise St	12th St		180	1	5	2	1	1	8	Easy

SCENARIO PHOTOS

Practice scenario one

Practice scenario two

142

Test scenario one

Test scenario two

Test scenario three

Test scenario four

Test scenario five

Test scenario six

Test scenario seven

Test scenario eight

Test scenario nine

Test scenario ten

APPENDIX I. EXPERIMENT TWO – INFORMED CONSENT

Letter of introduction with elements of consent: Software interface design

The purpose of this study is to examine software interface design on tablet PCs.

The experiment will be conducted in two sessions. During the study you may expect the following procedures to be followed. In the first session you will be asked to complete three cognition assessments and a short questionnaire. In the second session you will be asked to complete another short questionnaire, a training activity, and one exercise using a software interface.

There are no known risks to participation other than ergonomic concerns that are normally associated in the usage of the tablet PC.

There are no direct benefits to you as a participant other than the educational experience of being involved in an experiment. Your participation is helping the United States Census Bureau to develop better software.

By participating in this study, you will be provided with a $30 gift card. You will need to sign a receipt to receive this compensation.

Your participation in this study is completely voluntary and you may withdraw at any time. You may skip any part of this study that may make you feel uncomfortable.

Records identifying participants will be kept confidential. The information taken from this session will be assigned a unique code and will be used on forms instead of your name. Only researchers from Iowa State University working on this project will have access to data collected during this study. Study records will be kept confidential under password protected computer files. Data will be retained for two years and then will be destroyed.

The data collected in this research may be used for educational or scientific purposes and may be presented at scientific and/or educational meetings or published in professional journals. Published results will be in summary form only with no personal identifying information.

For further information about the study contact Michelle Rusch or Dr. Sarah Nusser at (515) 294-9773 or Dr. Les Miller at (515) 294-7588. If you have any questions about the rights of research subjects or research-related injury, please contact the IRB Administrator, (515) 294-4566, IRB@iastate.edu, or Director, Office of Research Assurances, (515) 294-3115, 1138 Pearson Hall, Ames, IA 50011.

Your signature indicates that you voluntarily agree to participate in this study, that the study has been explained to you, that you have been given the time to read the document and that your questions have been satisfactorily answered. You may receive a copy of the written informed consent upon request.

Participant's Name (printed) _____

_____   _____
(Participant's Signature)   (Date)

INVESTIGATOR STATEMENT

I certify that the participant has been given adequate time to read and learn about the study and all of their questions have been answered. It is my opinion that the participant understands the purpose, risks, benefits and the procedures that will be followed in this study and has voluntarily agreed to participate.

_____   _____
(Signature of Person Obtaining Informed Consent)   (Date)

APPENDIX J. EXPERIMENT TWO – TEST SCRIPT AND HARD

COPY TRAINING MATERIALS

Spatial ability and address matching study

Protocol test script

Iowa State University

June 16, 2008

21. Recruit participants and make appointments for them to be involved with the study
    a. Cognition testing appointments (June 11- 14)
    b. Software testing appointments (June 16 -21)

Cognitive testing

22. Run six sessions to administer three cognition tests (2-5 people per session)

    Prior to participant arrival for the first session, make sure to have the following:
    - A copy of the experimental protocol
    - Two copies of the letter of introduction with elements of consent for each participant
    - Participant assignment sheet
    - 10 copies of the VZ assessment
    - 10 copies of the RL assessment

23. When the participants have all arrived, welcome them into the conference room.

24. Introduce experiment and human subjects information

    Researcher: The purpose of this study is to investigate principles used to design software interfaces. Before we begin we need to provide you with more specifics about your rights as a participant, which are outlined in this document. Please read it carefully and sign it. Let me know if you have any questions.

25. Provide letter of informed consent and obtain signed consent.

26. Provide brief description of the format for the first and second session.

    Researcher: You are here tonight because you agreed to participate in a research study that investigates software interfaces. Today you will take three cognitive assessments,

complete a background questionnaire, and schedule a time to come back for a second appointment. During the second session you will test software that based on applications used by the United State Census Bureau. More details about the application will be provided during the second session.

27. Prior to assessments and questionnaire provide each subject with an ID number

28. Administer first assessment(s) and questionnaire
    Conference room
    a. Paper folding assessment (Part 1 and 2; 7 minutes)
    b. Reasoning assessment (15 minutes)
    Student computer area
    a. Administer background questionnaire

    Researcher: Now we have a short questionnaire for you to complete.

    Researcher enters id number on computer at a configuration screen for the background questionnaire and the software is initiated.

    Researcher: There are three screens to this questionnaire. After you have finished each screen, click the "next" button until you get the final screen. After you have completed the final screen there will be a button at the bottom of the screen that says "finish." Click on this "finish" button once you are finished and then we will move on to the next part of the experiment.

    b. Perspective taking assessment

29. Schedule second appointment

Software testing

30. Randomly assign one of the two treatments to participant

   Prior to participant arrival for the second session, make sure to have the following:
   - A copy of the experimental protocol
   - Participant assignment sheet
   - Two tablet PCs (one backup)
   - Debriefing survey- possibly on tablet?
   - Payment voucher
   - Compensation

31. Greet participant

   Researcher: Welcome back! We appreciate your continued participation in this study. I am going to be providing you with some instructions for our study. Please do not use the computer until I tell you it is time to do so.

   Today you will perform a series of exercises on the tablet PC. We will begin by describing what this exercise is. Then we will show you how to use the software to do the exercises. You will have a chance to practice the software on a couple of scenarios before you start the exercises.

   The task you will be doing today is based on a procedure called address listing. The U.S. Census Bureau uses this procedure to update their digital map files. In this procedure, a street map is displayed on a handheld computer. For each house on a block, the map shows a spot that marks its location. In our study, the map is displayed on the tablet computer, but is about the size of a handheld computer screen. The blue spot on the map marks the center of the land that a house sits on.

   Show the example visual to demonstrate a spot.

   If you worked for the Census Bureau, you would begin your work by checking addresses on a block. Imagine yourself walking along a city street. You would walk along the street and for each house on the block, you would check whether its map spot is displayed on the map and whether it is displayed in the correct location. By correct location, we mean that it is depicted on the correct side of the street and the neighboring houses are correct on the map. You would also make sure the map does not include extra houses that are not actually on the street.

   When we do this task in our experiment, we will show you photos of two sides of a street that you might see if you were actually outside walking along a block.

   Show photos that go with the map.

Suppose you are using this procedure and need to check on a specific target address. There are four possible outcomes.

1) *You find the target address on the block and on the map. The map spot depicting the address is also in the correct location on the map. In this case, the map is correct and you are done.*

2) *You find the target address on the block and the map, but the map spot is not in the correct location. This can happen if the map spot is displayed on the wrong side of the street or it does not have the correct neighbors. In this case, to correct the map, you would delete the existing map spot and put a new map spot in the correct place on the map.*

3) *You cannot find the address on the block and you cannot find it on the map. In this case, the map correctly reflects that no such house is actually on the block, so it does not need to be fixed.*

4) *You cannot find the address on the block, but a map spot for the address is shown on the map. This is an error on the map, and you would delete the map spot.*

In our study, we will ask you to check addresses by looking at the photographs on the monitors that represent the situation if you were walking along the block, and comparing this information with the map information.

The task that you perform follows the same basic steps.

1) *You will be given an address on the computer and asked whether you see this address on the photos on the monitors.*

2) *The computer will show you the map and ask you to find the address on the map.*

3) *If the address is on the map, you will be asked whether it is in the correct location.*

4) *If the map does not accurately reflect the target address, you will be asked to fix the map.*

Before we start the software, let's just review an example using these steps.

Addresses were placed on the map as the center point of existing parcels. A parcel is made up by the property lines for a house.

Show parcel and address spot visual.

Suppose our target address is 122 Trent St. You would begin by looking at the photos and trying to find 122 Trent St. Note that the streets are labeled with signs and houses have street numbers on them. See if you can find 122 Trent St.

So 122 Trent St. is not really present on the ground. Let's see if this address is on the map. Look for a spot on the map labeled with 122.
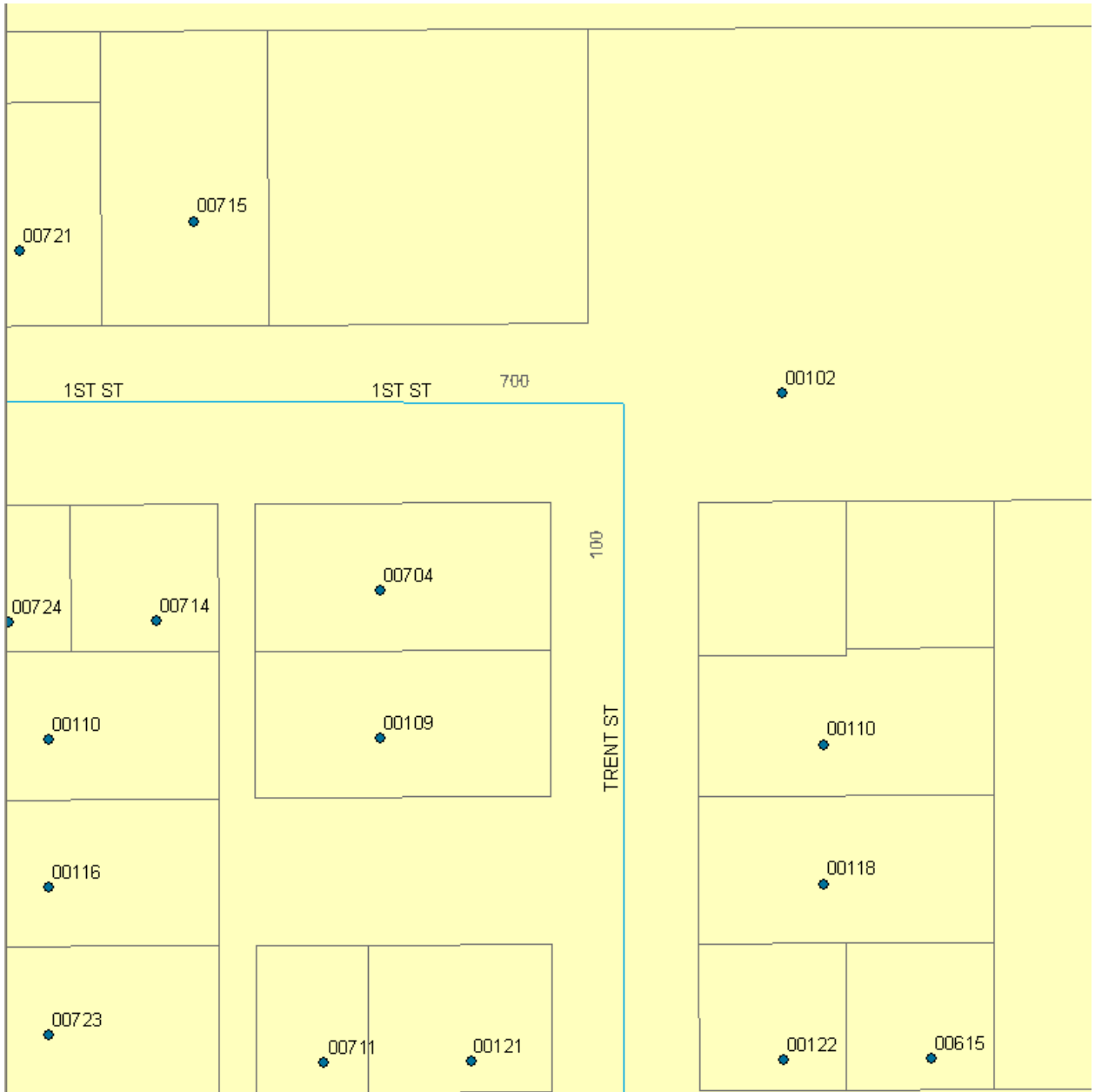
Remember, all of the spots on this map should represent housing units that are on the ground. You've found the address on the map, but it isn't in the photo. What do you think you should do?

In this case, you would proceed to delete the spot from the map.

Do you have any questions?

If not, then let's set up the computer for you.

32. Have both tablets set up to perform a calibration. Determine if the subject is left-handed or right-handed and adjust the tablets accordingly.

Researcher: Are you left or right handed?

Set stylus settings accordingly.

Researcher: Next, we will calibrate the screen. Please listen to my instructions before taking any action to calibrate the screen. At the upper right of the screen, you will see a cross. Each time you tap on this cross, it will move to a different corner of the screen. Use gentle and quick taps the cross each time it appears. Now please tap the first cross and continue to tap the other corners as the cross appears.

33. Look up the participant assignment list to obtain the condition under which the subject is to be run. This will be pre-determined through the random assignment prior to software use.

When the configuration screen appears, the researcher will type the assigned number for the particular subject. After the number has been typed they will then press the enter button. When the dialog window disappears and the application window appears everything is ready to go.

Participant introduced to practice scenario and trains using the specified button interface

Researcher: We will now teach you how to use the software to perform address and map checking tasks. Keep in mind that the focus of this study is not to test you, but rather to understand how well the software works. Also, please do not tap on the computer until I say to do so.

Once we start the exercise, the two monitors that are in front of you will show you two photographs that represent two sides of a street. In addition, the software will display a street address.

First we will step through two practice scenarios.

Go ahead and tap the start button in the center of the window on the screen and then wait for the next set of the instructions.

User taps start.

You see that after you tap on the start button that two photos appear on the monitors and the interface on the tablet is updated.

[For guided version only]

The yellow box includes a list of steps that you will go through for each of ten different scenarios. As you advance through each step the text will be bolded to show you where you what stage you are at in the scenario.

The white box to the right of the yellow one includes instructions specific to each step.

The software on the tablet computer identifies a street address that will be used in each scenario. We will call this the "target address." Below the yellow and white sections you see an address bar. The target address in this bar will change for each scenario that you go through.

[For unguided version only]
The software on the tablet computer identifies a street address that will be used in each scenario. We will call this the "target address." At the top of the display on the tablet you see an address bar. The target address in this bar will change for each scenario that you go through.

We are going to refer to the photos on the monitors as the "ground." One thing to keep in mind as you are comparing the photos on the monitors to the map is that the map will always be northbound, however the photos will not always be facing northbound. That is, they will change throughout the experiment and might be facing south, east, and westbound.

The first step in our procedure is to see whether the address can be found on the street images in front of you. Once you have determined whether the address is on the street photos, then answer the question in the box. Then wait for the next set of instructions.

The main exercise will involve comparing a map that will appear on the tablet to two real-world photos of streetscapes that will appear on the monitors. In this task, you will make changes to the map so it matches the real-world photos. Changes will include adding and deleting information. Additionally, there will be cases where no changes are required at all.

If incorrect:
As you can see, the computer has provided you with feedback that says your answer is not correct. Look again to see if you can find the address is in the photo.

Have subject try again, if s/he still fails to see it show the location in the photo

If correct:
You are providing answers that keep the application moving forward. In the case that you provided an answer that was not appropriate the computer will provide feedback that you need to try again. The next step involves finding the address on the map displayed by the tablet computer. You can see that the tablet screen now shows a map.

There are some tools that will help you work with the map. I will now explain these tools. Please wait until I have finished these instructions before tapping on the tablet again.

The first tool is the "zoom" tool. The slider at the bottom left of the screen will allow you to zoom in and out on the map. Zooming in will make the map show more detail in a smaller area. Zooming out will show more area, but with less detail. The + side represents more detail and the minus side represents less detail. To use this tool, you should first tap once on the map in the position which you would like to zoom upon. Then you can either tap on the plus or minus or on a specific level. A level represents a specific zooming magnification and each level is differentiated by the small tick marks below the zoom slider. To select a specific zoom level you can tap directly on the slider on the position in which you would like to zoom. You can also drag the zoom slider across multiple levels to get to your desired zoom level more quickly. As you zoom in closer you will be able to see the spots that we talked about at the beginning of the session. As you practice with these tools try to zoom in close enough to see these spots. Go ahead and practice with the zoom tools now.

The second tool is the "pan tool." The arrow buttons on the bottom right will allow you to reposition the map. If you want to see what is to the right of the map, push the arrow pointing to the right. If you want to see what is below the map, push the down arrow. The left and up arrows work in the same way. Go ahead and practice with the pan tools now.

The reset map button will reset the map to the original view. In the case that you are working on a scenario and "get lost" in a map you can use the reset map option to return to the original view. Now that you have become accustomed to the zoom and pan tools we will move forward with the practice exercise. Press the reset map button and then wait for the next set of instructions.

Your task now is to search for the target address using the pan and zoom tools so that you can see that address at a closer level. Go ahead and do that now. Keep in mind that this is practice. When you are done working with the map to find the address, let me know.

Participant indicates that they are done. If they have not zoomed in far enough or if the correct location is not within view provide guidance. Otherwise, proceed on to the next instruction.

The next step is to tell the computer that you are done looking for the target address. To do this, tap the 'Submit' button. You now see a new question on the screen. The question is asking if you were able to find the target address on the map. If the question happens to be covering up the map so that you cannot tell if the address is on the map you can tap the "Exit to map" button to make the question box disappear and then come back to it later.

Choose your response now.

If incorrect:
As you can see, the computer has provided you with feedback that says your answer is not correct. When you press the ok button the map will be reset to the original starting view. Try again to see if you can find the address spot is on the map. If the subject fails again, provide assistance.

If correct:
You are providing answers that keep the application moving forward. In the case that you provided an answer that was not appropriate the computer will provide feedback that you need to try again and the map would be reset. The next step is to determine if the spot for the address on the map is in the same location as it is in the photos. Remember that your goal is to fix the map so that it reflects what is in the photos. Compare the placement of the target address on the map to the house location in the photos. Do you see a mismatch?

If no or confused, provide guidance.

If yes:
As you can see, the map and the photos do not match. Your task now is to fix the map so that it is accurate. To make changes you can use the add and delete options. Look at the bottom left corner of the screen. You see the words add and delete with white circles next to each word. To use these options you must tap on the white circle and then tap on the map to use that tool.

To fix this scenario choose the delete option and then select the incorrect spot. Keep in mind that the tablet computer responds best to quick and gentle taps.

Screen should update with dialog that says, "Are you sure you want to delete address spot 123?"

Go ahead and tap ok.

Subject completes required change.

When making the delete change you provided answers that kept the application moving forward. In the case that you provided an answer that was not appropriate the computer would have provided feedback that you needed to try again and the map would have been reset.

Now tap submit.

34. Participant finalized the first practice scenario

Tap ok on the final screen of the first scenario.

You have just completed the first of two practice scenarios. You will now do one additional practice scenario before we start the actual exercise. In this second scenario, the direction of the streets in the photos will not be northbound. Keep that in mind as you are matching them to the map.

Press start to go ahead and begin.

Participant taps start.

The first step involves matching the target address to the photos. Look for the address on the photos and then answer the question.

Participant answers question.

After you answer the first question the next step is to use the zoom and pan tools to find the address on the map. Go ahead and do that now.

Participant has zoomed in.

Now that you have found the address location tap submit.

Participant taps submit.

You now see a question that asks you if the address spot is on the map. Answer that question now.

Participant answers question.

Compare the placement of the target address on the map to the house location in the photos. Do you see a mismatch?

If no:
Try again.
If they still fail or are confused provide guidance.

If yes, and correct:
The change that needs to be made for this scenario is similar to the last one with the exception of a couple new things. In this scenario there is spot on the map, however it is in the wrong location. This time you will add a new spot in addition to deleting the old one. To add, select the add option at the bottom of the screen and then tap where the new spot should be located. Use the photos as your guide.

35. *You have just completed the practice part of the study. You will now move on to the actual exercise. You will work through 10 scenarios that will be similar to the ones you just practiced. Keep in mind that there will be some differences. For example, you will not always be required to add and/or delete address spots. Do you have any questions?*

   Ok then, go ahead and get started. When you get to a screen that says "End of experiment" let me know and we will finish up our time together with a short questionnaire.

36. Participant executes the exercise

37. Participant completes the debriefing survey

38. Participant given incentive

   Researcher: Thank you for participating. I just need to get your signature on this payment voucher and then you will be free to go. If you know of any friends or co-workers that will also be participating in this study please wait to talk to them about what you did during the study until after they have completed the study themselves.

   After this, have them sign, hand them their compensation and send them on their way.

39. Participant excused to go

BIBLIOGRAPHY

Ackerman, P.L., Beier, M.E., & Boyle, M.O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology. General, 131*(4), 567–589.

Ackerman, P.L., & Kanfer, R. (1993). Integrating laboratory and field study for improving selection: Development of a battery for predicting Air Traffic Controller success. *Journal of Applied Psychology, 78,* 413-432.

Ackerman, P.L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information-processing. *Journal of Experimental Psychology: General, 11,* 288-318.

Ackerman, P.L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology, 77,* 598-614.

Ackerman, P.L., (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102,* 3-27.

Ackerman, P.L., (1988). Determinants of individual differences in skill learning: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117,* 288-318.

Ackerman, P.L., (1990). A correlational study of skill specifity: Learning, abilities, and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 883-901.

Ackerman, P.L., (1992). Predicting individual difference in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology, 77,* 598-614.

Allen, B. (1994). Cognitive abilities and information system usability. *Information Processing and Management, 30,* 177-191.

Arning, K., & Ziefle, M. (2006). Barriers of information access in small screen device applications: The relevance of user characteristics for a trans-generational design. In: User Interfaces for all. LNCS (2006), Berlin, Germany, Springer.

Atkinson, R.C., & Shiffrin, R.M. (1971). The control of short-term memory. *Scientific American, 225*, 2, 82-90.

Baddeley, A.D. (1986). Working Memory. Clarendon, UK: Oxford Univ. Press.

Baddeley, A.D., & Hitch, G. J. (1974). Working memory. In: G. H. Bower, Editor, *The psychology of learning and motivation Vol. 8*, (pp. 47–89) Academic Press, New York.

Benyon, D., Crerar, A., & Wilkinson, S. (2001). Individual differences and inclusive design. In: Stephanidis C. (ed.) User Interfaces for All: Concepts, Methods and Tools. LEA, Mahwah.

Berger, C.F., Lu, C.R., Belzer, S.J., & Voss, B.E. (1994). Research on the use of technology in science education. In Gabel, D. L. (Ed.), Handbook of Research on Science Teaching and Learning, Simon and Schuster, New York, pp. 466–490.

Booth, P., Fowler, C., & Macaulay (1987). An investigation into business information presentation at human-computer interfaces. In H. Bullinger & B. Shackel, Eds. *Interact '87*. Amsterdam: North-Holland, Elsevier, pp. 599-604.

Bryant, K.J. (1982). Personality correlates of sense of direction and geographical orientation. *Journal of Personality and Social Psychology, 43,* 1318-1324.

Bryant, D., & Tversky, B. (1999). Mental representations of perspective and spatial relations from diagrams and models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 137–156.

Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Science, 10,* 12, 551-557.

Büring, T., Gerken, J., & Reiterer, H. (2006). Usability of overview-supported zooming on small screens with regard to individual differences in spatial ability. In *Proceedings of the Working Conference on Advanced Visual interfaces* (Venezia, Italy, May 23 - 26, 2006). AVI '06. ACM Press, New York, NY.

Carroll, J.B. (1974). Psychometric tests as cognitive tasks: A new "structure of intellect." Princeton, N.J.: Educational Testing Service, Research Bulletin 74-16.

Carroll. J.B. (1993). Human cognitive abilities: A survey of factor-analytic studies. New York: Cambridge University Press.

Chen, C., & Rada, R. (1996). Interacting with hypertext: A meta-analysis of experimental studies. *Human-Computer Interaction, 11,* 2, 125-156.

Clarkson-Smith, L., & Halpern, D. (1983). Can age-related deficits in spatial memory be attenuated through the use of verbal coding? *Experimental Aging Research, 9,* 179-184.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence, 32*, 277-296.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671-684.

Dahlbäck, N., Höök, K., & Sjölinder, M. (1996). Spatial Cognition in the Mind and in the World: The case of hypermedia navigation. *The Eighteenth Annual Meeting of the Cognitive Science Society, CogSci'96*, University of California, San Diego, July.

Devlin, A.S., & Bernstein, J. (1995). Interactive wayfinding: Use of cues by men and women. *Journal of Environmental Psychology, 15,* 23-38.

Dillon, A., & Watson, C. (1996). User analysis in HCI- the historical lessons from individual differences research. *International Journal Human-Computer Studies, 45,* 619-637.

Egan, D. (1988). Individual differences in human-computer interaction. In: M. Helander (ed.), Handbook of Human-Computer Interaction, (pp. 543-568). Amsterdam: Elsevier Science Publishers.

Egan, D.E., & Gomez, L.M. (1985). Assaying, isolating and accommodating individual differences in learning a complex skill. In R. Dillon (Ed.), *Individual differences in cognition.* New York: Academic Press.

Ehret, B.D., (2000). Learning where to look: The acquisition of location knowledge in display-based interaction (Doctoral Dissertation, George Mason University, 2000). *Dissertation Abstracts International: Section B: the Sciences & Engineering, 60*(10-B), 5239.

Ehret, B.D. (2002). Learning where to look: Location learning in graphical user interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Changing Our World, Changing Ourselves* (Minneapolis, Minnesota, USA, April 20 - 25, 2002). CHI '02. ACM Press, New York, NY, 211-218.

Ekstrom, R.B., French, J.W., Harman, H.H., & Dermen, D. (1976). *Manual for Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.

Eliot, J., & Smith, I.M. (1983). An international directory of spatial tests. Windsor, England: NFER/Nelson; and Atlantic Highlands, NJ: Humanities Press.

Engle, R.W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19-23.

Fitts, P.M., and Posner, M.I., (1967). *Learning and skilled performance in human performance.* Belmont, CA: Brooks/Cole.

Golledge, R.G. (1991). Cognition of physical and built environments. In T. Garling and G. W. Evans (Eds.), Environment, cognition, and action: An integrated approach (pp. 35-62). New York: Oxford University Press.

Gomez, L.M., Egan, D.E. and Bowers, C., 1986. Learning to use a text editor: Some learner characteristics that predict success. *Human-Computer Interaction 2*, 1–23.

Goodman, J., Gray, P., Khammampad, K., & Brewster, S. (2004). Using landmarks to support older people in navigation. In S. Brewster & M. Dunlop (Eds.), Mobile human computer interaction, 38–48. Berlin, Germany: Springer.

Haigh, R. (1993). The aging process: A challenge for design. *Applied Ergonomics, 24,* 1, 9-14.

Hegarty, M., Richardson, A.E., Montello, D.R., Lovelace, K., & Subbiah, I.I. (2002). Development of a Self-Report Measure of Environmental Spatial Ability. *Intelligence 30,* 425-447.

Hegarty, M., Montello, D.R., Richardson, A.E., Ishikawa, T., and Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial layout learning. *Intelligence, 34,* 151–176.

Hintzman, D.L., O'Dell, C.S., & Arndt, D.R. (1981). Orientation in cognitive maps. *Cognitive Psychology, 13,* 277–299.

James, W. (1890). Principles of Psychology (Volume I). New York: Holt.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40,* 1, 1-17.

Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132*, 47-70.

Klatzky RL. (1997). Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections. *Conference on Raumkognition: Conference Proceeding.* Trier, Germany.

Kozhevnikov, M., Motes, M., Rasch, B., Blajenkova, O. (2006). Perspective-Taking vs. Mental Rotation Transformations and How They Predict Spatial Navigation Performance. *Applied Cognitive Psychology, 20,* 397–417.

Kozhevnikov, M., & Hegarty, M., (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory and Cognition, 29,* 745-756.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56,* 1479-1498.

Lohman, D.F. (1996). Spatial ability and g. In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement* (pp. 97-116). Mahwah, NJ: Lawrence Erlbaum Associates.

Mark, D.M., Freska, C., Hirtle, S.C., Lloyd, R., & Tversky, B. (1999). Cognitive models of geographical space. International Journal of Geographic Information Science, 13, 747-775.

McDonald, S. & Stevenson, R.J. (1998) Effects of text structure and prior knowledge on the learner on navigation in hypertext. *Human Factors, 40,* 18–28.

McGee, M.G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin, 86,* 889-918.

McGrath, D. (1992). Hypertext, CAI, paper, or program control: Do learners benefit from choices? *Journal of Research on Computing in Education, 24,* 4, 513-532.

Messick, S. (1976). *Individuality in learning.* San Francisco: Jossey-Bass.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology. General, 130*, 4, 621–640.

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. *Proc. ACM CHI'94 Conf.* (Boston, MA, April 24-28), 152-158.

Pak, R., Rogers, W. A., & Fisk, A. D. (2006). Spatial ability subfactors and their influences on a computer-based information search task. *Human Factors, 48,* 154-165.

Sein, M., Olfmann, L., Bostrom, R., & Davis, S. (1993). Visualization ability as a predictor of user learning success. *International Journal of Man – Machine Studies, 39*, 599 – 620.

Shepard, R.N., & Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology, 3,* 228-243.

Shepard, R.N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701-703.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127-190.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In H. W. Reese (Ed.), *Advances in Child Development and Behavior, 10*, 10-55, New York: Academic Press.

Thorndyke, P.W., & Goldin, S.E., (1981, March). Ability differences and cognitive mapping skill (N-1667-ARMY). Santa Monica, CA: Rand Corp.

Treisman, A. (1986). Features and objects in visual processing. *Scientific American, 255* (11), 114B-125.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97-136.

Tulving, E. (1985). How many memory systems are there? *Am. Psychol., 40,* 385–98.

Van der Veer, G.C., Tauber, M.J., Waern, Y., & Van Muylwijk, B. (1985). On the interaction between system and user characteristics, Behaviour & Information Technology, 4, 4, 289-308.

van Nimwegen, C. and van Oostendorp, H. (2007). Guidance in the interface and transfer of task performance. In *Proceedings of the 14th European Conference on Cognitive Ergonomics: invent! Explore!* (London, United Kingdom, August 28 - 31, 2007). ECCE '07, vol. 250. ACM, New York, NY, 225-232.

Vernon , P. E. (1950). *The structure of human abilities,* Methuen, London.

Vicente, K.J., and Williges, R.C., (1988). Accommodating individual differences in searching a hierarchical file system, *International Journal of man Machine Studies*, *29,* 647 – 668.

Vicente, K., Hayes, B.C., & Williges, R.C. (1987). Assaying and isolating individual differences in searching a hierarchical file system. *Human Factors, 29,* 349-359.

Wolfe, J.M., Cave, K.R., Franzel, S.L. (1989). Guided search: An alternative to the feature integration mode for visual search. *Journal of Experimental Psychology, 15,* 419-33.

Woods, D.D. (1984). Visual momentum: a concept to improve the cognitive coupling of person and computer. *Int. J. Man-Mach. Stud. 21,* 3 (Sep. 1984), 229-244.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science, 21*, 2, 179-217.

Zhang, J., & Norman, D.A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18*, 87-122.

Ziefle, M., & Bay, S. (2006). How to overcome disorientation in mobile phone menus: a comparison of two different types of navigation aids. *Human-Computer Interaction, 21*, 4, 393-433.

ACKNOWLEDGEMENTS