Navigating Through Buildings with Impaired Vision:
Challenges and Solutions

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Amy Ashwin Kalia

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisers:
Gordon E. Legge Paul R. Schrater

June 2009

Acknowledgements

This achievement would not have been possible without the support of so many people. I would especially like to thank my advisers, Gordon E. Legge and Paul R. Schrater. I am very grateful to have worked with two brilliant mentors who not only challenged me, but who also had the patience, enthusiasm, and confidence to see me through the most difficult parts of this journey.

I also thank Daniel J. Kersten and Gary W. Oehlert for serving on my committee, and for always being helpful over the years.

I am grateful to the Low-Vision Lab family for both academic and social support. I particularly want to thank Deyue Yu and Miyoung Kwon for their continuous help, encouragement, and friendship. Thanks also to Chris Kallie for being a source of inspiration, a teacher, and a friend.

Also, I would like to thank Nicholas Giudice, Rudrava Roy, and Advait Ogale for collaborations on the research in Chapters 1 and 2. Thanks also to the undergraduate assistants who helped with data collection and analysis: Joann Wang, Julie Yang, Ameara Aly-youssef, Ryan Solinsky, and Jennifer Mamrosh. I also thank Paul Beckmann and Brian Stankiewicz for creating the Photorealistic virtual environments described in Chapter 2, and Chris Kallie, Advait Ogale, and Sandeep Dhull for the design and construction of the LED apparatus described in Chapter 4.

I am grateful for all my friends, new and old, for being there to share the best and worst moments of the past six years.

Finally, I thank my family, especially my mom, dad, and brother. They are my foundation, and I am forever grateful for their love and faith in me.

Funding for this research was provided by: The University of Minnesota College of Liberal Arts and Department of Psychology's Graduate Research Partnership Program, and NIH grants T32 HD007151 (Interdisciplinary Training Program in Cognitive Science), T32 EY07133 (Training Program in Visual Neuroscience), EY02857, U.S. Department of Education (H133A011903) through a subcontract through SenderoGroup LLC., EY017835-01 (Designing Visually Accessible Spaces), and EY015616-03 (Magnetic Wayfinding).

Dedication

To my parents, Rita and Ashwin Kalia.

Abstract

Navigation is the ability to plan and follow routes between locations, often with an internal or external map of the environment. Navigation is complex because it relies on mechanisms from several brain subsystems, including perception, movement, and memory. Vision is an important way to access environmental information for navigation. Consequently, independent navigation is a significant challenge for individuals with visual impairment. This thesis describes three studies that investigate how real or simulated visual impairment affects the ability to navigate inside buildings. Furthermore, these experiments explore methods for compensating for the loss of visual information, either by using other senses or by using assistive technology.

The visual information in an environment that is useful for navigation can be categorized as two types: geometric (visual information conveying layout geometry such as hallways and intersections), and non-geometric (features other than geometry such as lighting, texture, and object landmarks). The first experiment (Chapter 2) describes the effects of visual impairment and age on the use of these two types of visual information for navigation. Participants (older and younger individuals with normal or low vision) learned novel building layouts in two types of computer-generated virtual displays, one displaying geometric layout features only, and the other displaying geometric features and also non-geometric features. Participants were tested on their knowledge of the layout by generating a map of the layout and by finding specific locations in the corresponding real environment. We found that age rather than visual impairment influenced the reliance on the two types of visual information; older individuals demonstrated more accurate layout knowledge after learning with non-geometric

information compared to learning with geometric information. Younger individuals were able to use both types of information to learn layouts.

In the second experiment (Chapter 3), visually-impaired individuals were tested on their ability to follow verbal route instructions provided by an indoor navigation technology. The instructions described distances in one of three ways: feet, number of steps, and travel time in seconds. The route-finding technology improved the ability of participants to find rooms in an unfamiliar building, measured by the distance they travelled between rooms and the number of times they asked for help. Performance was most improved when distances were conveyed in steps, and this was also the preferred distance mode as reflected by participant ratings.

Humans typically can determine their location in an environment using two sources of information: visual landmarks and path integration (movement information obtained while walking). The third experiment (Chapter 4) investigated how two factors, congruence and reliability, influence the integration of visual and walking information for localization in a hallway. We predicted that humans integrate information: 1) only when they perceive themselves to be near a landmark after walking (congruency), and 2) by weighing each information source according to its reliability. Normally-sighted participants judged their location in a hallway after viewing a target and then walking blindfolded to either the visual target or to a slightly different location. Participants viewed targets in two conditions that manipulated visual reliability: normal viewing and blurry viewing. As predicted, participants integrated both sources of information only when they perceived the visual and walked locations to be the same. Also, when information was integrated, they sometimes relied more on walking information to determine their location as the reliability of the visual information decreased with blur. This experiment tested and confirmed a statistical model of human perception in a novel domain. The results suggest perceptual mechanisms for adapting to visual loss in navigation tasks.

Together, these three studies enhance our understanding of the effects of visual impairment on navigation ability. These studies also suggest that information provided by other senses or assistive technology can improve navigation ability with low vision.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Navigation is the ability to plan and follow paths from one location to another. The ease and frequency with which we navigate obscures the complexity of this process. Successful navigation requires the seamless integration of several brain subsystems, including perception, movement, and memory, thus making it a challenging topic of study. Research from a diverse array of perspectives has contributed knowledge about the underlying workings of navigation, including animal (Gallistel, 1990) and human (Golledge, 1999) behavioral studies, computer science and robotics (Thrun, 2002) and neuroscience (Redish, 1999). The understanding of navigation provides insight into the more general question of how the brain uses perceptual information to accomplish goal-oriented behavior.

Background

In 1948, Tolman published a set of experiments describing the behavior of rats in several types of maze navigation tasks. The results illustrated that the rats were behaving in ways that suggested a more complex understanding of the environment than could be explained by simple stimulus-response mechanisms. For example, the rats could make novel shortcuts to the location of the feeder even if that particular route had never been associated with a reward. These observations spurred the idea that the animals established and maintained a mental representation of the environment, called the cognitive map, when navigating. Questions regarding the nature of the cognitive map have fueled much debate and investigation in animal and human navigation. While some suggest that cognitive maps preserve geometric relationships (Gallistel, 1990), there are numerous examples of how geometrical properties are distorted in human memory (Tversky, 1993). Alternative theories describe cognitive maps as representations of route and topological information (Kuipers, Tecuci, & Stankiewicz, 2003; Siegel and White, 1975), or views of the environment associated with actions (Mallot & Gillner, 2000; Schölkopf & Mallot, 1995).

Vision provides abundant information that is useful for the acquisition and storage of spatial relations. The available visual features in an environment can be categorized into two types: geometrical or non-geometrical. Geometric visual information conveys the structure of the environment, for example the hallways and intersections in a building. Non-geometric visual information refers to features other than geometry, such as lighting, texture, and objects serving as landmarks. Research on both animals and humans suggests that geometric information is preferentially used for navigation, although non-geometric information is used to differentiate locations that are geometrically similar (see review in Chapter 2). The study described in Chapter 2 explores the use of geometric versus non-geometric visual information for three groups: young adults, older adults, and visually-impaired adults. We found that older adults, whether or not they are visually-impaired, rely more on non-geometric information for navigation.

How well is navigation accomplished when visual information is unavailable? Imagine trying to walk from your office to the mailroom while blindfolded. In addition to the risk of colliding with obstacles, it is difficult to navigate without visual cues to indicate the distance and direction to important locations, such as landmarks, intersections or the destination. Independent navigation is a daily challenge for individuals with visual impairment and it significantly impacts their quality of life (Golledge, Marston, Loomis, & Klatzky, 2004). Despite the attention given to mobility and obstacle avoidance (Kuyk & Elliott, 1999; Marron & Bailey, 1982; Turano, et al., 2004; West, et al., 2002), less is known about the impact of visual impairment on navigation.

Given the advantages of visual information for navigation, an important question is whether people without vision can develop accurate and functional spatial representations (review by Thinus-Blanc & Gaunet, 1997). Several studies have found that people with visual impairment since early in life do not acquire spatial representations of familiar environments to the same level of accuracy as people with normal vision (Rieser, et al., 1992; Ungar, Blades & Spencer, 1996). However, blind people fare better when missing visual information is compensated for by non-visual information. For example, blind individuals are able to accurately follow routes and update their location with respect to locations conveyed by sound or spatial language (e.g. “2 o’clock, 16 feet”) (Loomis, Golledge & Klatzky, 1998; Loomis, Lippa, Klatzky, & Golledge, 2002). In an effort to aid navigation by individuals with visual impairment, recent technologies have capitalized on the ability to develop spatial representations from non-visual modalities (reviews by Giudice and Legge, 2008, and Loomis, Golledge, & Klatzky, 2001). Chapter 3 describes an effort to convey route information via verbal descriptions of layout geometry to visually-impaired pedestrians.

Path integration, used when landmarks are not available (Gallistel, 1990), is a method of navigation that involves continuously tracking the direction and distance traveled from a starting location. Animals and humans are able to use visual (optic flow), proprioceptive, and vestibular information to accurately estimate angular and linear displacements during movement (Loomis, Klatzky, Golledge & Philbeck, 1999). Yet, noise in these sensory signals result in position estimates that are imperfect and these errors accumulate over time. Therefore, it is beneficial for navigators to use a combination of landmark and path integration information. The interaction between landmarks and path integration has been explored in both the animal (Collett, Graham, & Durier, 2003) and human (Ellard & Shaughnessy, 2003; Nardini, Jones, Bedford & Braddick, 2008) literature by testing which strategy is employed when they give conflicting information. In the study described in Chapter 4, we use statistical models of human perception to determine if information from these two strategies are combined optimally, based on measures of reliability of the corresponding sources of information.

The overall goal of this thesis is to understand how the loss of visual information affects the ability to navigate, and how this information can be replaced or compensated for. In Chapters 2 and 3, we study how technology can be used to compensate for reduced visual information, and in Chapter 4 we examine how the perceptual system might compensate by relying more on non-visual sensory information.

Overview of Experiments

Humans with normal vision are able to navigate effectively with only geometric information, although additional visual features can improve performance in some cases. In Chapter 2, we investigated the types of features that are useful for navigation by older people and by people with visual impairment. We studied the effects of visual impairment and age in conjunction because people with visual impairment are typically older. The practical impetus for this study was to determine if people with visual impairment could use virtual displays to learn the layouts of buildings; the results might be used to design assistive navigation technologies for the visually-impaired.

We tested three groups of participants: younger adults, older adults, and individuals with visual impairment with a broad age range. Participants learned four building layouts, two in different types of virtual environments (VEs), one with a map displayed on a computer and one by walking around in the real space. One virtual environment displayed only geometric information in high contrast (Sparse VE), and the other displayed additional realistic visual features (Photorealistic VE) (Figure 2.2). After learning the layouts in one of these four conditions, participants were tested on their layout knowledge by constructing maps of the floors and by navigating to specified locations in the corresponding real floor. Interestingly, we found that individuals with visual impairment performed similarly to their age-matched counterparts; they were able to navigate accurately in both the Sparse and Photorealistic VEs. Yet, older people performed better after learning in the Photorealistic VEs compared to the Sparse VEs, suggesting that they relied more on non-geometric features to learn the layouts. Therefore, age, more so than visual impairment, influences reliance on non-geometric features for learning layouts. This result suggests that visually-impaired individuals with some functional vision can use visual displays for navigation.

Whereas the study in Chapter 2 explored the type of information needed to develop a map-like representation of the environment, Chapter 3 focused on the informational requirements of following routes. We explored how to best convey route information within the context of developing indoor navigation technology for visually-impaired users. Two critical issues that developers must address when creating navigation technology for the blind are: what information needs to be conveyed, and what is the best way of conveying this information using a non-visual sensory modality, such as audition.

Previous research indicates that blind users can learn environments using verbal descriptions of layout geometry (Giudice, 2004; Giudice, Bakdash & Legge, 2007). The study described in Chapter 3 tested the ability of people with visual impairment to follow verbal route instructions provided by a technology, called the Building Navigator, developed in the laboratory of Gordon Legge at the University of Minnesota.

The Route-finding feature of the Building Navigator computes the shortest path between the starting and goal locations, and produces a series of instructions describing the route. The instructions consist of the direction and distance to several intersections along the route where the user makes a turn. Distances are described in one of three modes: feet, number of steps, and time to travel in seconds. The distances in steps and seconds were individually calibrated according to participants' step length and walking speed. We tested the ability of blindfolded sighted and visually-impaired participants to follow the route instructions using the three distance modes. We also compared the ability to find rooms with and without the technology. One experimenter guided the participant while another experimenter measured the time taken to follow the route, the path travelled, and the number of bystander queries. The results indicated that blindfolded sighted participants performed significantly better with the technology than without, and distance in feet and steps improved performance in all measures. Visually-impaired participants travelled shorter paths and made fewer bystander queries with the technology compared to without. Performance was also best with distance in steps. Furthermore, in post-experiment surveys, visually-impaired participants indicated that they preferred distance in steps compared to the other conditions. These outcomes reveal that verbal instructions comprising geometrical descriptions of the environment allow for accurate route-following. Also, the preference for distances in steps by visually-impaired users is instructive for the design of future technologies.

Chapter 4 addressed the question of how the perceptual system might adapt to visual impairment. Although it is assumed that people with visual impairment rely more on non-visual information for travel, clinical observations indicate these individuals often rely more on their residual vision than they should, often leading to hazardous collisions (Ludt & Goodrich, 2002). Within the context of navigation, people with normal and impaired vision can employ landmark and path integration to utilize different sensory modalities to localize themselves in an environment. When using visual landmarks, the observer's representation of location consists of his or her direction and distance from a viewed object (e.g. "I am 3 feet behind the water fountain."). The distance and direction to objects are obtained from a variety of visual cues, including location in the visual field and depth information. When an object or goal location is unavailable (e.g. it is too far or obstructed from view), humans can use path integration to determine their location while moving, which requires integrating movement information over time. An observer can determine changes in movement from vestibular and kinesthetic sensory information.

The goal of the study described in Chapter 4 was to determine how humans combine visual and walking information to estimate their location in a building environment. We addressed this question by drawing from previous research in human perception examining how sensory information is combined to make perceptual judgments. These studies provide evidence that humans combine sensory information in a statistically optimal fashion by weighting each source of information according to its reliability. The reliability of a sensory cue is inversely related to the variability of its estimates. Accordingly, we predicted that information obtained by viewing (visual landmark strategy) versus walking (path integration) to objects would be weighted according to the reliability of the corresponding sensory information to determine the observer's location. We tested this prediction by manipulating the reliability of visual information; observer's viewed objects either with normal (high reliability) or blurry (low reliability) vision. We hypothesized that observer's would weight walking information more when visual information was blurry compared to when vision was normal.

We tested thirteen normally-sighted participants in a fifteen meter hallway under normal lighting conditions. First, we measured the reliability of visual and walking information. Participants viewed targets (high-intensity LEDs on the floor) with normal and blurry vision, or walked to targets while blindfolded. The distances of the targets ranged from 5 to 11 meters. Participants then judged either the location of the viewed target or their own location after walking by marking the position on a tactile map. As expected, the reliability of visual information decreased with blur, i.e., the variability in the estimated target location increased. Accordingly, we predicted that when both visual and walking information were available, the weighting of walking information would increase as the reliability of visual information decreased. These predictions were tested using a cue conflict paradigm in which participants judged their location in a hallway when the visual and walking information provided discrepant information. Participants first viewed a target with normal or blurry vision, and then walked blindfolded until stopped by an auditory signal. The walked distance was the same as the visual target or conflicted by +/- 0.25, 0.5, or 0.75 meters. Participants marked on the tactile map the location they walked to. Then, they judged whether the visual and walked locations were the same or different. The results of the cue conflict task showed that for a target distance of 9 m, participants increased their reliance on walking information when vision was blurry, as predicted by the theory of optimal cue combination. For targets at 7 m, the results did not clearly demonstrate the extra reliance on walking information. These mixed results provide cautious support for optimal cue combination under some circumstances, but imply the role of additional factors affecting how sensory information is weighted. One possibility is that since information is presented sequentially, instead of simultaneously like in previous studies of simpler perceptual tasks, the order of presentation may affect weighting.

The results of this study also have implications for navigation with visual impairment. An open question is whether the perceptual systems of these individuals re-weight walking information as visual information becomes less reliable to optimally estimate the traveler's location. Perhaps the undue reliance on vision described by Ludt and Goodrich (2002) is related to an inappropriately greater weighting of visual information.

Together, these studies provide insight into how degraded visual information can be supplemented, either perceptually or with technology, for successful navigation. The fact that participants, either with actual or simulated visual impairment, can incorporate alternate sources of information to navigate speaks to the adaptive nature of the human perceptual system. The results of these studies will contribute to the development of technological and rehabilitative methods for improving everyday functioning of people with visual impairment.

Chapter 2:

Learning Building Layouts with Non-geometric Visual Information: The Effects of Visual Impairment and Age

This chapter has been published as: Kalia, A.A., Legge, G.E., & Giudice, N.A. (2008). Learning Building Layouts with Non-geometric Visual Information: The Effects of Visual Impairment and Age. *Perception, 37(11)*, 1677-1699

Introduction

A typical building contains abundant visual features for aiding navigation, from geometric cues about the structural layout of the floor plan to cues unrelated to the layout geometry such as the presence of objects (e.g., pictures, water fountains) and image characteristics (e.g., textures, color and lighting). The current study addresses how two important participant characteristics, visual impairment and age, influence the types of visual information needed for developing an accurate mental representation of a novel virtual environment (VE). First, we ask whether rendering of purely geometrical information is sufficient for navigation in virtual buildings with visual impairment, and whether the addition of non-geometric visual features helps or hinders. Second, because the prevalence of visual impairment is much higher in old age, we ask whether age influences the use of geometric and non-geometric visual information.

Geometric and Non-geometric Cues

In this study, geometric cues refer to the spatial configuration of hallways, specifically their length and intersection connectivity. In Figure 2.1A, the geometric features are the hallways extending to the left, right, ahead, and behind. Non-geometric visual features are distinct from layout geometry, and in Figure 2.1A include the bulletin board with postings, the trash cans in the corridor, and the lighting patterns on the walls and floor.

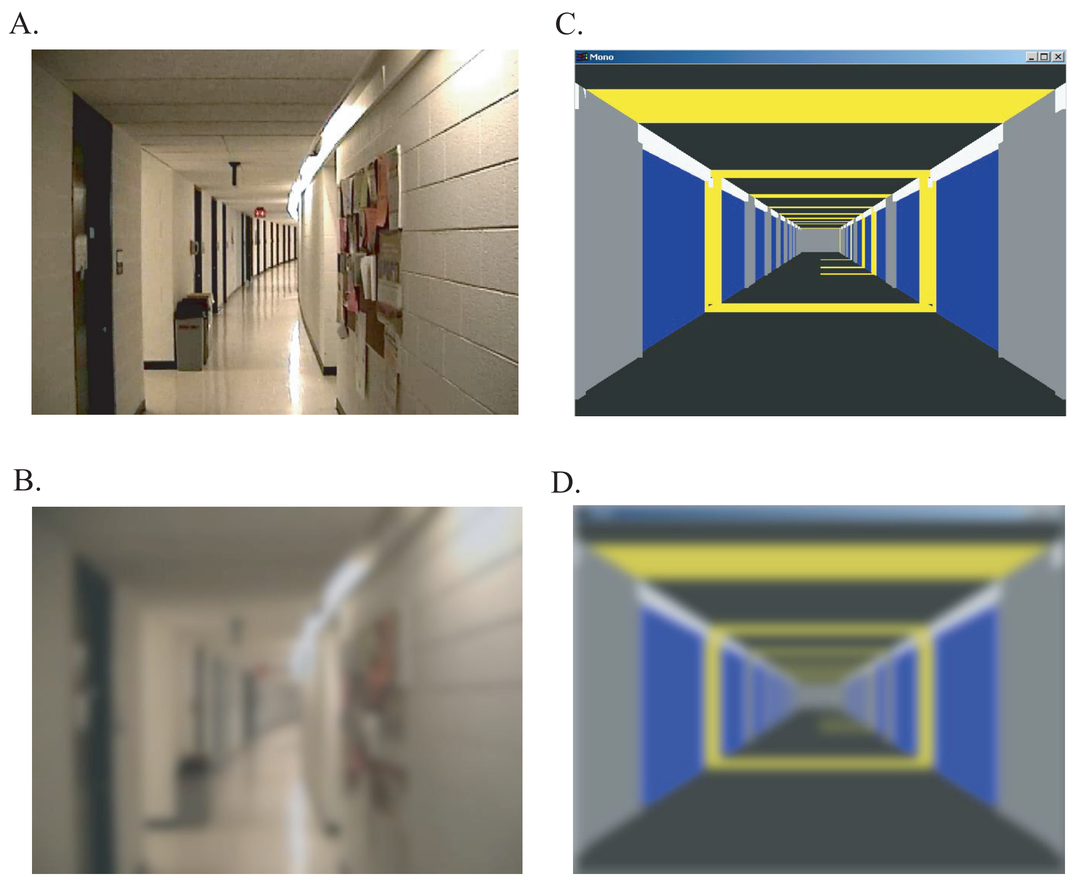

Figure 2.1. Example of a four-way hallway intersection as viewed normally (A) and with simulated visual impairment produced with blurring and contrast reduction (B). The geometric and non-geometric features depicted in A become less distinct in B with simulated visual impairment. In a virtual environment, geometric features can be displayed in high-contrast colors (C). Under reduced visual conditions, branching hallways are easier to detect in the virtual environment (D) compared to the real environment. Although these images do not necessarily simulate the subjective experience of people with low vision, they give some idea of the reduction of visual information associated with reduced spatial resolution or reduced contrast sensitivity.

Previous work on both animal and human spatial cognition suggests that information about layout geometry is preferentially encoded when learning a space. After exploring a rectangular box, rats look for the target the same percentage of time at the correct corner as at the geometrically-equivalent opposite corner, despite the presence of unique non-geometric cues (Cheng, 1986). Furthermore, pre-verbal human children tend to use geometric information to locate the position of a toy in a rectangular room even when wall color, a non-geometric feature, provides more specific information (Hermer & Spelke, 1996). These findings support the notion of a “geometric module” in the brain dedicated to using information about the relative position of surfaces in an environment to compute orientation (Gallistel, 1990).

Non-geometric cues are useful for learning environments when geometric cues are ambiguous. Monkeys and other species will rely more on non-geometric cues, such as cards with distinctive patterns, when the information they provide about target location conflicts with geometric cues (Gouteux & Thinus-Blanc, 2001; Kelly, Spetch, & Heth, 1998; Sovrano, Bisazza, & Vallortigara, 2002; Vallortigara, Zanforlin, & Pasti, 1990; see review by Cheng & Newcombe, 2005). However, these species are still able to use geometric information for localization when non-geometric information is not available.

There is also evidence that humans are biased towards using geometric information when navigating through more complex spaces, such as the inside of a building. Non-geometric information, such as large objects placed at various intersections, improves navigation efficiency (measured by the route-distance traveled) (Lessels & Ruddle, 2005; Ruddle, Payne, & Jones, 1997), and is useful for locating specific rooms and remembering where to make turns (Ruddle, et al., 1997). However, non-geometric information does not result in better knowledge of overall layout configuration (Ruddle, et al., 1997). Furthermore, participants demonstrate more accurate knowledge of geometric compared to non-geometric information, even during the first few exposures to a novel indoor environment (Stankiewicz & Kalia, 2007). These studies suggest that non-geometric cues provide some advantages when navigating through an environment, but they are not necessary for developing an accurate mental representation of layout information.

The current studies explored whether people with varying degrees of visual abilities and ages demonstrate similar use of non-geometric visual information when learning unfamiliar, large-scale layouts. Although previous studies suggest that younger, normally-sighted individuals do not rely on non-geometric cues to develop an accurate mental representation of a layout, it is not known whether the same is true for older adults or people with visual impairments. We tested this question by comparing learning of layouts in two types of virtual environments, one that displayed only geometric features (Sparse VE) and another that displayed both geometric and non-geometric features (Photorealistic VE). Like the studies by Lessels & Ruddle (2004, 2005), we did not select a single type of non-geometric feature to include in the virtual environments, because it was unclear which visual features humans choose to use in real spaces. Instead, the Photorealistic VE allowed participants to use the range of non-geometric information available in real environments.

Low Vision Navigation

The term "low vision" refers to any chronic visual impairment that affects everyday functioning and is not correctable by glasses or contact lenses. There are two distinct problems associated with navigating with low vision: obstacle avoidance and wayfinding. Much attention has been given to the problem of how specific visual impairments make it difficult to detect and avoid obstructions along a path (Kuyk & Elliott, 1999; Marron & Bailey, 1982; Turano, et al., 2004; West, et al., 2002). The current study focuses on wayfinding behavior in people with low vision, including their ability to learn unfamiliar layouts and to use this information to plan and execute paths between specific locations.

Low vision may influence wayfinding and the building of accurate mental spatial representations in two ways. First, it can be more challenging to visually extract layout geometry, making it more difficult to navigate between locations. If so, wayfinding should be improved by enhancing the saliency of geometric visual features in a layout. Virtual environments can accomplish this by, for example, depicting hallways and intersections in high-contrast colors. In Figure 2.1, the intersecting hallways are more salient in the virtual environment (2.1D) than in the real environment (2.1B) under conditions of blur and reduced contrast. Secondly, people with low vision may rely less on non-geometric information if reduced acuity, reduced field or low contrast sensitivity makes it difficult to resolve and identify these features. If such objects are hard to recognize, they might even act as distracting clutter which interferes with the extraction of cues to the geometrical layout. Thus, we might expect little or no improvement in wayfinding when non-geometric features are present in an environment. Based on these predictions, we hypothesized that low-vision participants would learn layouts better in the Sparse VE, which displayed only high-contrast geometric information, compared to the Photorealistic VE, which displayed additional non-geometric information.

This study also addressed the practical question of whether low-vision individuals can use enhanced computer displays to learn a layout of a building prior to visiting the real space. Pre-journey learning, using maps on a digital touchpad augmented with auditory cues, has proven useful for blind individuals when navigating in the real environment (Holmes, Jansson & Jansson, 1996). The current study investigated whether the same is true for visual representations displayed on a computer, such as maps or first-person virtual environments.

Aging and Navigation

The leading causes of vision impairment in the U.S. are due to age-related eye diseases, such as age-related macular degeneration, cataract, diabetic retinopathy and glaucoma (Eye Diseases Prevalence Research Group, 2004). It is estimated that in the year 2000 there were approximately 3.3 million individuals older than 40 with visual disabilities, with prevalence growing significantly with age (Eye Disease Prevalence Research Group, 2004). It is likely that age and visual impairment interact in their effects on spatial navigation. Consequently, our study also explores how age influences the use of geometric and non-geometric visual information when learning novel indoor layouts.

Several studies have shown that navigating in novel layouts becomes more difficult with age. A recent study by Sjölinder, Hook, Nilsson, and Andersson (2005) had younger (mean age = 25.5) and older (mean age = 66.9) individuals navigate through a virtual grocery store by simulating the real-world task of searching for items on a shopping list. The virtual rendering included visual details such as store shelves with different products and textured walls and floors. The older participants spent more time and were less efficient at finding the grocery items in the virtual store compared to the younger participants. They also developed less accurate survey knowledge of the store layout, and particularly overestimated distances between locations. Another study demonstrated that older (> 65.1 years) participants take more time and exhibit more errors when learning a route in a visually-rich virtual environment compared to younger participants (Moffat, Zonderman, & Resnick, 2001). This is despite controlling for other factors such as computer experience and gender. Animal research on the use of place cells to encode spatial information also indicates that older rats have difficulty developing spatial representations for new environments and associating target locations with visual cues (Rosenzweig, Redish, McNaughton, & Barnes, 2003; Wilson, et al., 2004).

Age may also affect the ability to use geometric visual information during spatial learning. In a study by Moffatt & Resnick (2002), younger (25-45 years), middle-aged (45-65 years), and older (65-93 years) adults were trained to locate a hidden platform by virtually swimming in a Morris Water Maze. The circular pool, viewed from a first-person perspective on a desktop computer, was surrounded by walls with an irregular shape as well as distinct objects that served as non-geometric cues. When asked to draw a map of the environment, older participants were less accurate at depicting the outer wall but were able to reproduce the non-geometric features of the environment. Furthermore, the older participants had difficulty locating the target with respect to the geometric cues provided by the shape of the walls compared to younger participants, but no differences between age groups were found when non-geometric cues were available. Older adults also exhibit more error than younger adults when using a geometric representation of a route (a map consisting only of layout geometry) to navigate inside a building (Wilkniss, et al., 1997). These results suggest that older individuals may rely more on non-geometric rather than geometric cues when learning locations within a space. Accordingly, in the current study we hypothesize that older adults will have a less accurate representation of layouts learned in the Sparse VE, which only depicts layout geometry, compared to younger adults, but both age groups will perform similarly in the Photorealistic VE.

Current Study

To summarize, the current study addressed whether non-geometric visual information usefully supplements (or, in the case of low vision, obscures) geometric information for learning novel indoor layouts. We tested the effects of two participant characteristics on the use of geometric and non-geometric information: 1) visual impairment, and 2) age. Participants learned layouts in two types of environments displayed on a desktop computer, a Sparse VE depicting only information about layout geometry (hallway length and intersections), and a Photorealistic VE displaying a full range of visual features, such as posters on the walls, color and lighting patterns, in addition to geometric information. For comparison, participants also learned layouts using two common methods, a Map and by exploring a Real building.

We measured two aspects of learning, the rate of acquisition of layout information and the accuracy of the resulting mental representation. The rate of acquisition was measured by how much exploration was needed to learn the locations of several targets. Target localization and map drawing were used to measure the accuracy of the resulting mental representation and whether it contained the minimal information required to travel between locations (route-based knowledge), or additional information about the overall configuration of the layout (survey knowledge). Experiment 1 tested young adults with normal vision, Experiment 2 tested a heterogeneous group of low-vision individuals, and Experiment 3 tested older people with normal vision.

Experiment 1

This experiment tested how well young, normally-sighted individuals learned layouts in four presentation modes: two virtual environments (Sparse and Photorealistic) displayed from a first-person viewpoint, a Map, and a Real building. The Sparse and Photorealistic VEs were used to manipulate the non-geometric visual information available when learning layouts. The main goal of this experiment was to replicate previous findings that non-geometric visual information does not improve the accuracy of mental representations of space, but may aid other aspects of navigation for younger normally-sighted individuals.

Two control conditions were included in this experiment. The Map condition assessed learning when global information about the layout was available. The Real building condition provided an ecologically valid control for comparison, and also indicated whether non-visual cues are crucial for learning layouts. In this paper, non-visual cues refer to proprioceptive and vestibular information. Previous studies suggest that non-visual information allows for more accurate spatial updating as measured by judgments of the direction to objects (Chance, Gaunet, Beall, & Loomis, 1998; Waller, Loomis, & Haun, 2004) or to a starting location (Klatzky, et al., 1998). Also, non-visual information increases the efficiency of searching for targets in both a visually-sparse and a photorealistic VE (Ruddle & Lessels, 2006). By testing participants in these four conditions, we compared how acquisition and accuracy of layout knowledge is effected by the information available about the layout during learning.

Method

Participants

Sixteen undergraduate students (mean age = 19, SD = 1, 8 males, 8 females) from an introductory psychology course participated in the experiment. Eleven of the 16 participants were located and surveyed four years after the study on their video game experience. Six of the eleven did not have experience with video games at the time of the study; the other five spent 0.5 to 3 hours per week playing video games that involved navigating through virtual environments. Participation was voluntary and was compensated with extra credit or monetary payment. All participants had normal or corrected-to- normal visual acuity and had little to no exposure to the building layouts tested in the experiment. Participants also provided informed consent.

Materials

The layouts used for testing were obtained from several topologically-distinct floors in the psychology building at the University of Minnesota. Computer representations of these layouts were created by mapping them onto a grid of nodes connected by line segments (see Figure 2.2C as an example). Each node represented a possible position in the computer-based virtual environment and only discrete moves between nodes were allowed. Each line segment represented a hallway unit connecting one node to another, which corresponded to an approximate distance of 4.6 m (15 feet) in the real layout. Thus, for both virtual environments and the Map displayed on the computer, participants moved from one node to another, or the equivalent of 4.6 m, every time they made a forward key press.

Each layout contained four target locations represented acoustically but not visually. We have no reason to believe that memory for auditory versus visual targets is different since previous work has shown that spatial representations acquired visually or verbally are functionally similar (review by Loomis, Klatzky, Avraamides, Lippa & Golledge, 2007). When participants reached a target location, a computer voice stated the name of the target (e.g. “Target Cat”). The speech output was adjusted beforehand for all participants to ensure that it was highly intelligible. The targets were located either at dead ends or intersections and were chosen to be spread out across the floor. This required participants to explore most of the layout in order to find all target locations.

The same keystroke interface was used to move in all three computer-simulated environments (Sparse VE, Photorealistic VE, and Map conditions). Participants used key presses on the number keypad to either translate forward one hallway segment (key “8”), rotate 90 degrees to the right (“4”) or rotate 90 degrees to the left (“6”). In the Map condition, the keystrokes moved a cursor that provided a visual indicator of the participant’s current location (see below).

The Map and Sparse environments were viewed on PCs with 18 inch (45.7 cm) screens (ViewSonic E90f and MultiSync FE950+) and the Photorealistic environments were viewed on a Macintosh with a 17-inch (43.2 cm) screen (Apple Studio Display). Different computers, and therefore screens, were used because of the software requirements of the virtual environment programs. The room lights remained on during testing because it was more comfortable for visually-impaired (Experiment 2) and older (Experiment 3) participants and made the experiment less intimidating.

Sparse VE

In the Sparse condition, participants were presented with a virtual representation of the real floor viewed from a first-person perspective. These sparse virtual layouts were created on a rectangular grid and only rendered the hallways and intersections of the floor; they did not include non-geometric features.

To simplify the rendering, curved hallways were approximated as straight hallways. It is likely that modest curvature is a feature that humans do not typically preserve in their cognitive maps of a space. People typically encode environments as having a grid-like structure, with 90 degree intersections, even if the real environment deviates from a grid structure (Tversky, 1981). It remains possible that even if hallway curvature is not represented geometrically, it could be encoded as a non-geometric image feature based on effects of shading, contour or occlusion associated with the curvature.

Hallways were rendered with contrasting blue and yellow color patterns on the walls and ceiling; additional yellow markings on the wall and floor indicated the presence of intersecting hallways (Figure 2.2A). Virtual movement from one node to the next was continuous and included optic flow cues, thereby simulating real movement. Translations and rotations took about 1.3 seconds. These environments were generated using Virtual Reality Modeling Language (VRML), and were displayed using Virtual Reality Utility (VRUT; University of California, Santa Barbara).

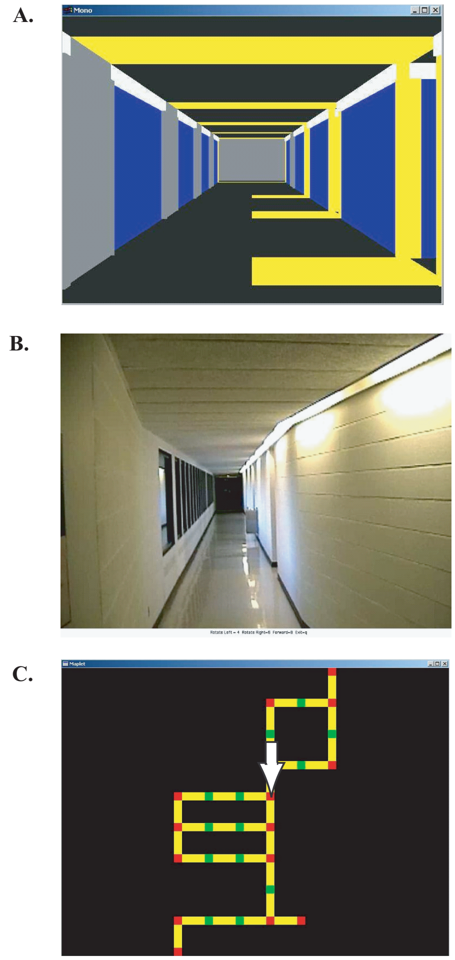

Figure 2.2. Examples of the same layout rendered in the A) Sparse VE, B) Photorealistic VE, and C) Map presentation modes. The views in A and B are from the same location and orientation as indicated by the white arrow in C (the arrow is enlarged for this illustration). In A, the yellow markings on the right side of the floor and wall indicate a hallway going to the right at the next three intersections.

Photorealistic VE

The virtual environments in this condition were also rendered from a first-person perspective, but they included both the geometric and non-geometric cues available in the real environment (Figure 2.2B). The environments were created using movies of real space recorded by mounting a camera on a robot arm attached to a moving cart. The eye height of the camera was approximately 1.5 meters (5 feet) off the ground. For rotations, the arm was turned at a controlled pace, taking approximately 1.7 seconds for a 90° rotation. For translations, the cart was pushed a distance of about 4.6 m in 7.3 seconds. The movie was broken into clips simulating every possible forward move between nodes and every 90-degree rotation at intersections. Key presses generated the corresponding movie of the translation or turn.

Translational and rotational speed in the Photorealistic VE was slower than the Sparse VE because of the physical limitation of moving the cart. The difference in movement speeds in the two VEs did not affect participants' perception of distance, as measured by a distance estimation task not reported in this paper.

Map

Participants learned one layout using a map displayed on a computer screen (Figure 2.2C). The map displayed nodes connected by yellow segments and each segment represented 4.6 m in the real layout. The nodes were green if they were located within a continuing hallway and red if they were at a dead end or an intersection. The display also included a white arrow indicating the participant's current position and heading. Participants explored the layout by moving the arrow with key presses that indicated movement with respect to the arrow. For example, if the arrow was facing down on the screen, pressing "4" (the left turn key) would turn the arrow left with respect to its initial direction, meaning the arrow would point towards the right side of the screen. The reason for making movements with respect to the arrow, the participant's projected location, was to be consistent with exploration in the Sparse and Photorealistic VEs in which key presses correspond to egocentric movements from a first-person perspective.

Procedure

Participants were tested in a within-subjects design by learning a different layout in all four conditions. They went through the procedure a total of five times, once for practice at the beginning of the experiment and then once for each of the four conditions. The order of the conditions as well as the layout-condition pairings were counterbalanced across participants with the restriction that only two of the four layouts could be learned in the Photorealistic condition (due to availability of materials). The layouts were quite different with the intention that learning would not transfer.

First, participants practiced moving in each type of environment using key presses and were told that each forward move was equivalent to moving fifteen feet (4.6 m) in the real building. Then they learned a practice layout and performed tests assessing layout knowledge described below.

Acquisition of Layout Knowledge

Each of the four experimental conditions started with an exploration period of a novel layout. Participants were explicitly told to explore until they definitively knew the locations of the targets and were familiar with the layout, but perfect knowledge of layout topology was not required. The exploration period ended when participants indicated they knew the target locations or they reached a maximum number of forward moves (determined by the size of the layout). Participants then performed a learning test that required them to navigate to each of the four target locations within a maximum number of forward moves (twice the number of moves needed to take the shortest path) in the same presentation mode as during exploration. The purpose of the learning test was to have participants achieve a common level of learning before they were assessed on their knowledge of the layouts. If participants did not pass this criterion level of learning, they resumed exploration. Participants were allowed to explore the layout for a maximum of four iterations to pass the learning test.

The rate of acquisition of layout knowledge was measured by recording the total number of forward moves used by the participant to explore the layout before passing the learning criterion. The acquisition score was calculated as the proportion of the number of moves used to the total number allowed in four possible exploration periods. For example, if the total number of moves allowed, summed over four exploration periods, was 200, and the participant used a total of 150 moves, the acquisition score would be 150/200 = 0.75. A low score indicated that layout information was attained with little exploration, whereas a score of 1.0 indicated that participants used all the exploration time allowed by the experimenter.

Assessment of Layout Knowledge

Participants performed two tasks to test the accuracy of the knowledge acquired during the exploration session. Participants first made maps of the layouts, which assessed whether they had a survey understanding of the environment. The second task was to find routes between target locations in the real space; this assessed transfer of layout knowledge to the real building.

Map Drawing. Participants were given a 16 x 16 grid of dots, and were asked to draw a map of the layout by connecting the dots with lines representing fifteen-foot (4.6 m) hallway segments. They also labeled the positions of the targets.

The accuracy of the map drawings was determined using a method adapted from Waller, Knapp, & Hunt (2001). This method only assessed the accuracy of target placement on the maps relative to each other and thus did not consider how accurately the corridor network was depicted.

Error (*E*) was the sum of the distances, measured in hallway units, of the estimated locations of targets from their actual locations (Equation 2.1 from Waller, et al. (2001)). The vectors ξ and ψ were the coordinates of the actual locations of targets for each map and vectors *x'* and *y'* were the coordinates of the estimated locations of targets on the corresponding participant maps. The coordinates for actual and estimated target locations were centered relative to the average of these coordinates (Equation 2.2 adapted from Waller, et al. (2001)). The participant target placements were also rotated and scaled to produce a minimum error score for each map.

$$E = \sum_{i=1}^{n} \sqrt{(\xi_i - x_i')^2 + (\psi_i - y_i')^2} \qquad \text{Equation 2.1}$$

$$x' = x_i - \left(\frac{\sum_{j=1}^{n} x_j}{n} \right) \quad \text{and} \quad y' = y_i - \left(\frac{\sum_{j=1}^{n} y_j}{n} \right) \qquad \text{Equation 2.2}$$

Target Localization. Participants were instructed to find the locations of targets in the real building layout corresponding to the virtual layout they had explored. They navigated to targets in an order determined by the experimenter. If participants incorrectly localized a target, they were taken to the correct location to start the next trial. This method prevented errors made on a single trial from affecting performance on subsequent trials. Acoustic information about target locations was not available during this test. Participants received a score for target localization accuracy for each of the four layouts learned. This score was calculated as the proportion of targets correctly located out of the total number of targets (four per layout).

Data Analysis

Learning was assessed by measuring both the acquisition rate and the accuracy of spatial knowledge when participants explored layouts in the four presentation modes. Repeated measures analyses of variance (ANOVA) were conducted for acquisition rate and map error measures. Friedman tests (the nonparametric version of the Repeated Measures ANOVA) were conducted for target localization performance since the data for this measure were non-normal and could not be sufficiently corrected with transformations. Bonferroni-corrected pairwise tests determined which presentation modes accounted for significant differences in performance. Effect size, or how much the variance in the data was accounted for by presentation mode, was measured by partial eta squared (η_p^2).

Results

Acquisition of Layout Knowledge

There was a significant effect of Presentation Mode on the rate of acquisition of layout knowledge ($F(3, 45) = 28.785$, $p < 0.001$, $\eta_p^2 = 0.66$). As shown in Figure 2.3, acquisition was significantly faster (i.e., fewer moves were required to reach criterion learning) in the Photorealistic VE than in the Sparse VE ($p = 0.003$). The Map and Real building conditions also required significantly less exploration compared to the Sparse VE (Map and Real building: $p < 0.001$). Three participants included in this analysis did not pass the learning criterion in the Sparse VE condition, but the accuracy of their layout knowledge was still assessed. These data are included in the following analyses because excluding them did not alter the pattern of results.

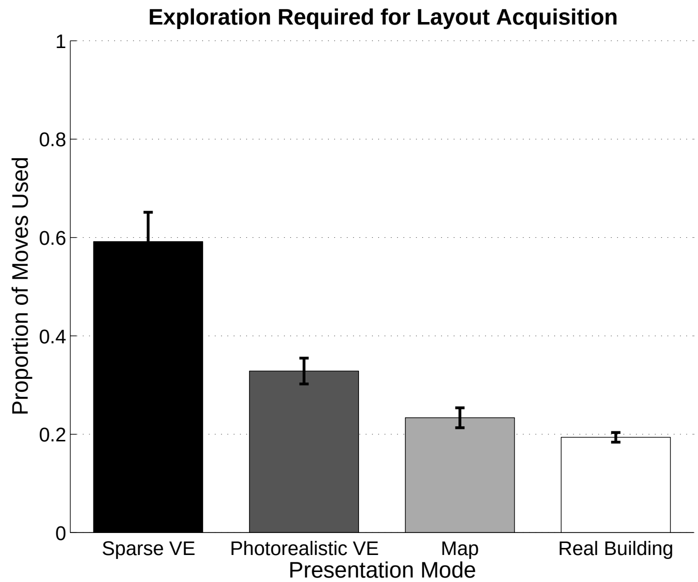

Figure 2.3. Mean proportion of moves used by younger normally-sighted participants to reach criterion learning in four presentation modes. Error bars indicate +/- 1 standard error. The proportion of moves used is calculated as the number of moves used divided by the total number of moves allowed to reach criterion learning.

Accuracy of Layout Knowledge

No significant difference was found in target localization performance between the four presentation modes (η_p^2 = 0.09) (Figure 2.4). Although a significant difference was found in map drawing error (F(3, 45) = 4.497, p = 0.008, η_p^2 = 0.23), pairwise comparisons were not significant (Figure 2.5). Individual scores on the target localization and map drawing tasks averaged across conditions were significantly correlated (p < 0.001) with a correlation coefficient of -0.838. The overall trend for both tasks was that performance was least accurate in the Sparse VE condition. Interestingly, a high proportion of targets were correctly located in the real space after learning in all presentation modes, indicating that the visual information provided in each condition, even the Sparse VE, was sufficient for learning to navigate between targets.

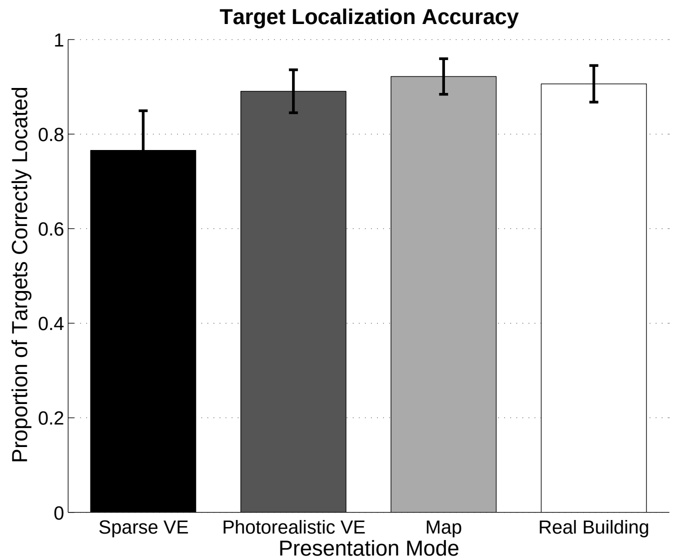

Figure 2.4. Mean proportion of targets correctly localized by younger normally-sighted participants in four presentation modes. Error bars indicate +/- 1 standard error.

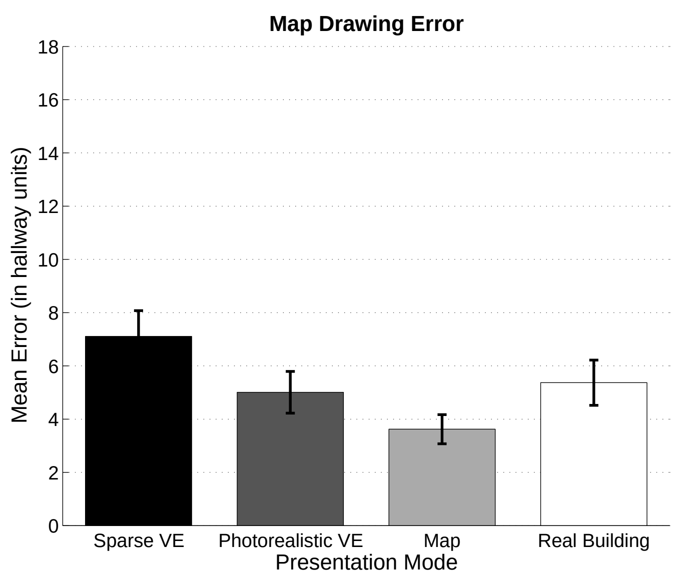

Figure 2.5. Mean error (measured in number of 15-foot hallway units) in the map drawing task by younger normally-sighted participants in four presentation modes. Error bars indicate +/- 1 standard error.

The Effect of Video Game Experience on Performance

Considering the age range of these participants, it is possible that video game experience influenced performance in the virtual environments. Games such as first-person action and driving/racing games require navigating from a first-person perspective through virtual environments that are visually sparse, at least at the time of testing (ca., 2004). Participants who played these types of games may have exhibited better performance in the Sparse VE compared to people without video game experience. Therefore, we conducted a post-experiment survey to evaluate the video game experience of our participants four years after they were tested. We were able to obtain data for only eleven of the sixteen participants. Six of these participants did not have experience with video games at the time of the study; the other five spent at least half an hour per week playing video games that involved navigating through virtual environments.

The target localization performance and map drawing error of the video game players and non-video game players are shown in Figures 2.6 and 2.7. Wilcoxon Rank Sum tests (the nonparametric version of the independent samples t-test) did not reveal significant differences across conditions between participants with and without video game experience in either the target localization or map drawing tasks. We acknowledge a trend that video game players have higher target localization accuracy in all conditions compared to non-video game players, as shown in Figure 2.6, but this difference was not significant. Because we were specifically interested in whether video game experience accounted for similar performance in the Sparse and Photorealistic VEs, we conducted a Wilcoxon Signed Rank Test (the nonparametric version of the matched-samples t-test) to compare performance in the VEs of the non-video game players. We did not find significant differences in either target localization accuracy or map drawing error, suggesting that video game experience did not selectively improve performance in the Sparse VE.

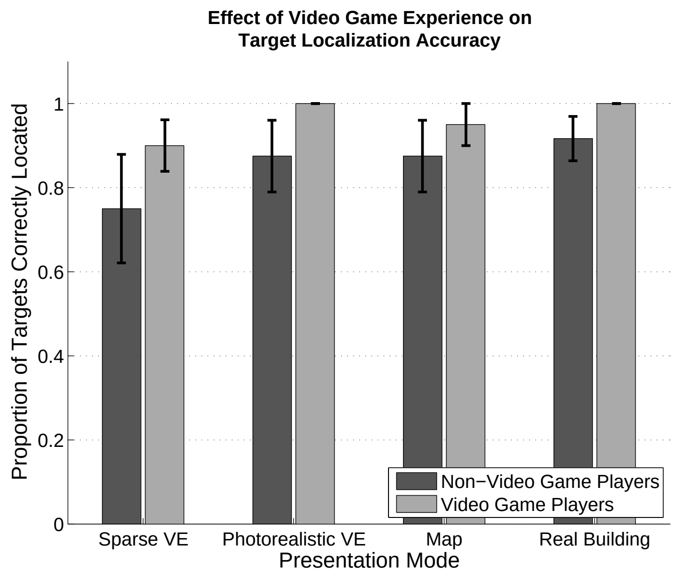

Figure 2.6. Mean proportion of targets correctly localized by younger normally-sighted participants with and without video game experience. Error bars indicate +/- 1 standard error.

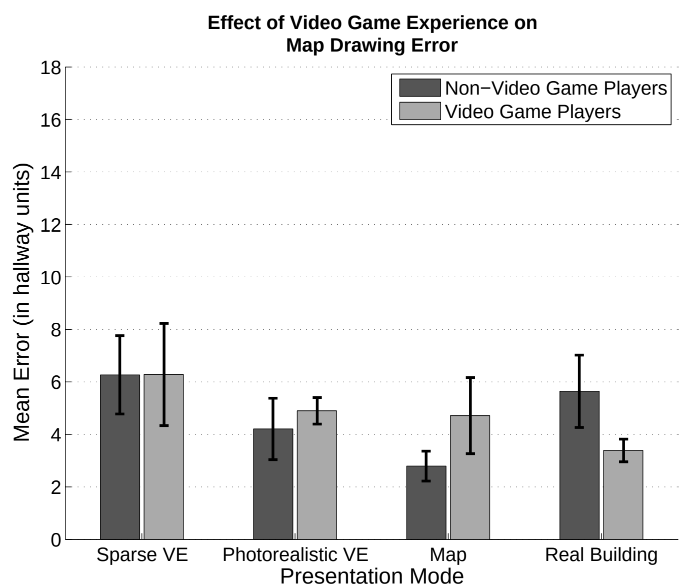

Figure 2.7. Mean error (measured in number of 15-foot hallway units) in the map drawing task by younger normally-sighted participants with and without video game experience. Error bars indicate +/- 1 standard error.

Discussion

For the normally-sighted young adults tested in this experiment, the type of visual information rendered in the first-person virtual environments significantly affected the acquisition rate but not the accuracy of layout knowledge. Non-geometric cues accelerated the acquisition of layout information; significantly less exploration was needed in the Photorealistic VE compared to the Sparse VE to learn the locations of targets in the layout. However, participants demonstrated similar layout knowledge after learning in either the Sparse or Photorealistic VE, which supports previous findings that non-geometric cues do not improve the accuracy of the acquired mental representation (Ruddle, et al., 1997; Thompson, et al., 2004). We interpret these results as showing that non-geometric information speeds up acquisition of layout knowledge for novel environments, but is not necessary for developing an accurate mental representation of the space.

Learning layouts in the Map and Real building conditions did not significantly improve performance compared to learning in the Photorealistic VE, indicating no advantage of having global layout information or non-visual cues. It is possible that the localization task was not challenging enough to distinguish performance between the Photorealistic VE, Map, and Real building conditions. Also, the types of VEs or tasks used to measure performance may influence these results. Previous studies in photorealistic VEs found that non-visual cues increase the efficiency of navigational search (Ruddle & Lessels, 2006) and the ability to accurately point to targets in a layout (Waller, et al., 2004). Non-visual information also aids spatial updating when only sparse visual information is available (Chance, et al., 1998; Klatzky, et al., 1998). Like the current study, Ruddle & Peruch (2004) found that non-visual information did not improve target localization, measured as the distance travelled to targets, in mazes that included non-geometric cues. Our study is in agreement with this finding that non-visual information may not be useful for learning the locations of targets when non-geometric visual information is available.

Video game experience did not account for comparable layout knowledge after learning in the Sparse and Photorealistic VEs. It is known that action video games improve perceptual abilities such as the capacity, spatial distribution, and temporal characteristics of visual attention (Green & Bavelier, 2003), but their influence on spatial navigation abilities needs further investigation. Studies that have examined the effect of prior computer experience and attitudes towards computers on virtual environment learning have found mixed results (Waller, 2000; Waller, et al., 2001); therefore, it is still unclear how computer experience might influence performance.

Experiment 1 replicated previous findings that non-geometric information does not improve the accuracy of mental representations of building layouts for younger, normally-sighted individuals. The next experiment investigated whether the same is true for people with visual impairments.

Experiment 2

In Experiment 2, we predicted that low-vision participants would acquire more accurate cognitive maps when learning in the Sparse VE compared to the Photorealistic VE because critical geometric information was rendered with high-contrast features, and extraneous non-geometric cues were removed. The presence of non-geometric features could hinder learning if they are hard to identify or are treated as visual clutter rather than useful information. The idea of contrast enhancement for low vision is incorporated into closed-circuit TV (CCTV) magnifiers (Lund & Watson, 1997) and has been explored in image-enhancement algorithms for face recognition (Peli, Lee, Trempe, & Buzney, 1994) and TV images (Peli, 2005).

A practical goal of this experiment was to assess the potential utility of virtual visual displays as navigation aids for low vision. Blind individuals can use tactile maps (Holmes, et al., 1996) and verbal descriptions of layout geometry (Giudice, 2004; Giudice, Bakdash, & Legge, 2007) to learn a layout before visiting it. Sparse VEs that display high-contrast renderings of geometrical information may also be useful low-vision aids if they improve wayfinding.

Method

Experiment 2 followed the same procedure as Experiment 1 except for the alterations discussed below.

Participants

Thirteen people with low vision (mean age = 41, SD = 18, 6 males, 7 females) were recruited from the community. Participants had a wide range of visual characteristics as described in Table 2.1. Visual acuity was measured using the Lighthouse Distance Visual Acuity Chart. Contrast sensitivity was measured with the Pelli-Robson Chart. Diagnosis and field status were obtained from reports supplied by the participant's ophthalmologist or optometrist. The participants were not familiar with the psychology building where testing occurred and received monetary compensation for their time.

We were non-selective in the nature of the visual conditions of the participants because our focus was on the general effect of low vision rather than specific types of visual impairment. Enrollment was restricted to individuals with no known cognitive deficits or physical deficits limiting mobility, and to individuals with vision adequate enough to perceive the features in the virtual environments necessary for the navigation tasks. The experiments were time consuming, requiring five to six hours for each participant. Because of the length and demands of the testing paradigm, our sampling of low-vision participants was skewed towards younger individuals with long-standing forms of visual impairment.

Participant	Age	Gender	Diagnosis	LogMAR Acuity	Log Contrast Sensitivity	Field Loss
1	25	F	Aniridia	0.78	1.65	None
2	20	M	Retinitis pigmentosa	1.34	1.35	Peripheral
3	29	F	Albinism	0.9	1.5	Peripheral
4	31	M	Retinitis pigmentosa	0.58	0.9	Peripheral
5	52	F	Retinopathy of prematurity, nystagmus	0.96	1.05	Peripheral
6	24	F	Retinopathy of prematurity	1.18	0.9	None
7	38	M	Stargardts disease	1.04	1.05	Central
8	53	F	Optic atrophy, no vision in left eye	1.52	0.3	Details not available
9	18	M	Scotoma caused by brain tumor	1.02	1.2	Details not available
10	67	M	Macular degeneration in left eye, cataracts in both	0.88	0.9	Central
11	51	M	Macular degeneration (Doynes' Macular Dystrophy)	1.04	1.65	Central
12	66	F	Cone-rod dystrophy	0.70	1.20	Central
13	62	F	Retinitis pigmentosa	1.24	0.45	Peripheral with spread to central

Table 2.1. Description of thirteen low-vision participants tested in Experiment 2. Central field loss is scotoma within 5 degrees of the fovea. Peripheral field loss is scotoma anywhere outside 5 degrees from the fovea.

Materials

The environments displayed on a computer were the same as in Experiment 1 with one exception. In the Map condition, the cursor was an enlarged, blinking, white and black triangle embedded in a gray square. The size of the triangle could be altered for each participant to ensure that its position and pointing direction could be seen.

Procedure

Participants were trained in each presentation mode displayed on the computer. The experimenter described in detail the features displayed on the computer screen (e.g., doors, windows, and lighting reflections in the Photorealistic VE) and verified that participants could see and describe the geometry of nearby intersections as rendered on the screen. Next, participants practiced moving through the layout using key presses, and followed a series of directions given by the experimenter (e.g. “Turn at the next intersection”). Participants had to successfully follow the experimenter’s directions before continuing with testing.

Map Drawing

Low-vision participants used Legos to create a map of each layout. They were instructed to build hallways by connecting Lego pieces on a 15 x 15 grid of nodes. Each Lego piece represented one fifteen-foot (4.6 m) hallway segment in the real environment. Participants also pointed to the target locations on their Lego map.

<u>Results</u>

Acquisition of Layout Knowledge

The rate of acquisition was significantly affected by Presentation Mode ($F(3, 36) = 19.107$, $p < 0.001$, $\eta_p^2 = 0.61$); participants spent significantly more time exploring in the Sparse VE than either with the Map or in the Real building ($p < 0.001$) (Figure 2.8). However, there was not a significant difference in exploration time between the Sparse and Photorealistic VEs. Five of the eleven participants used as many or more moves to explore the Photorealistic VEs compared to the Sparse VEs. Therefore, acquisition of layout information was not necessarily easier with non-geometric information than with only geometric information, but varied by individual.

Six participants (from Table 1 participants 5, 7, 8, 10, 12 and 13) did not pass the learning test in the Sparse VE, and of them, two (8 and 10) did not pass criterion in the Photorealistic VE. Also, there was a significant order effect ($F(3, 36) = 4.847$, $p = 0.006$, $\eta_p^2 = 0.29$), but pairwise comparisons did not reveal significant differences between conditions.

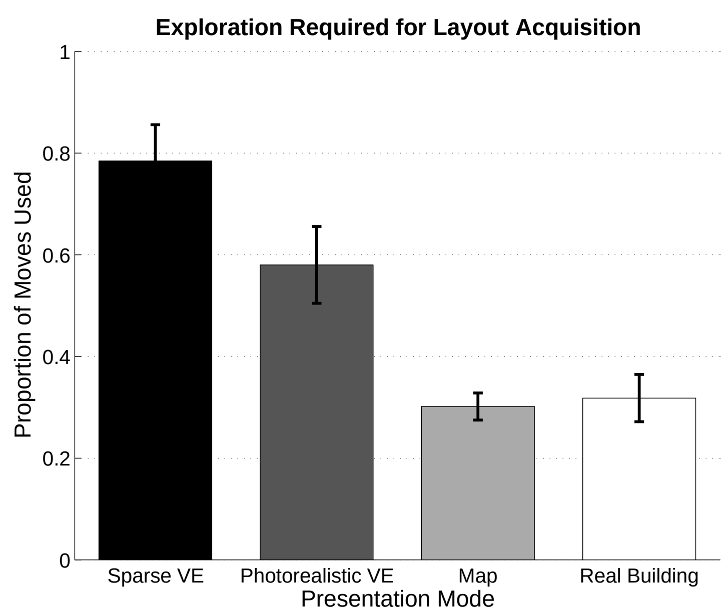

Figure 2.8. Mean proportion of moves used by low-vision participants to reach criterion learning in four presentation modes. Error bars indicate +/- 1 standard error.

Accuracy of Layout Knowledge

Significant differences were found in both target localization accuracy ($\chi^2(3)$ = 18.73, $p < 0.001$, $\eta_p^2 = 0.48$) and map drawing error ($F(3, 33) = 9.503$, $p < 0.001$, η_p^2 = 0.46). These results are presented in Figures 2.9 and 2.10. Performance in both tasks after learning in the Photorealistic VE was better compared to the Sparse VE ($p < 0.01$), but similar to learning in the Map and Real building conditions. These findings suggest that, contrary to our prediction, low-vision participants developed more accurate mental representations of layouts in the Photorealistic VE compared to the Sparse VE. Furthermore, target localization was less accurate in the Sparse VE condition compared to the Map ($p = 0.002$) and Real building ($p < 0.001$) conditions. Map drawing error was also greater after learning in the Sparse VE compared to learning in the Map condition ($p = 0.007$).

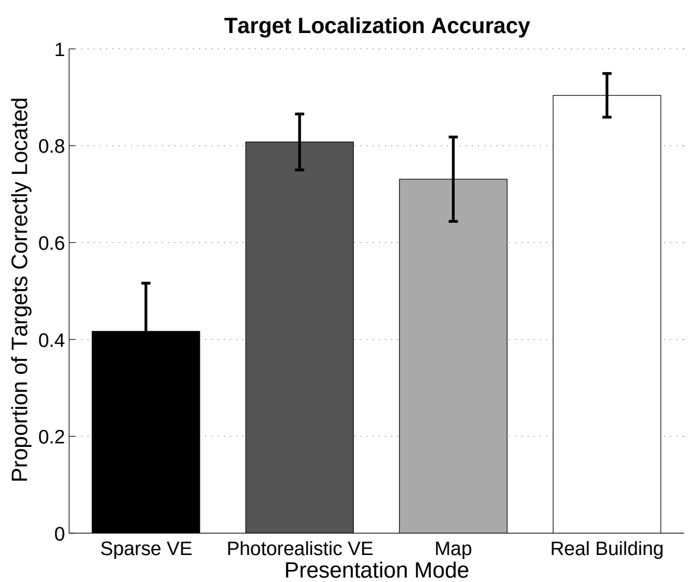

Figure 2.9. Mean proportion of targets correctly localized by low-vision participants in four presentation modes. Error bars indicate +/- 1 standard error.

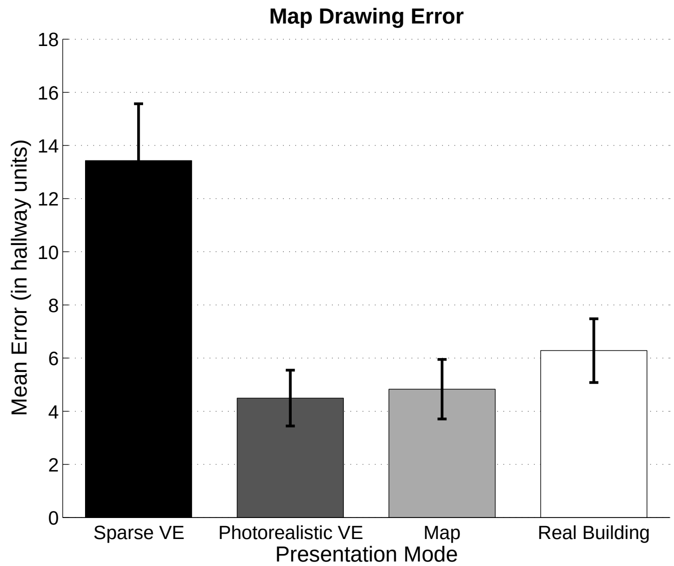

Figure 2.10. Mean error (in hallway units) in the map drawing task by low-vision participants in four presentation modes. Error bars indicate +/- 1 standard error.

Relationship between Performance and Ocular Factors

Learning by participants in the virtual environments may be related to the nature of their visual impairment. Although we did not select participants with the goal of linking their performance to ocular factors, we did examine the relevant correlations to determine what relationships might exist. Also, because we recruited participants who had enough vision to see the VEs, there is likely a restriction in the range of acuity, contrast sensitivity, and field loss of those tested.

Acuity did not correlate with any of the performance measures for either the Sparse or Photorealistic VE. Strong correlations were found between the proportion of moves used during learning and log threshold contrast sensitivity in both the Sparse VE ($r = -0.696$, $p = 0.008$) and the Photorealistic VE ($r = -0.700$, $p = 0.008$). Previous research has also shown that contrast sensitivity has a greater impact on orientation and mobility than acuity (Marron & Bailey, 1982). We did not have sufficient information on visual-field status to compute correlations, but previous studies indicate that visual field loss is associated with increased errors in obstacle avoidance (Turano, et al., 2004). Field loss from eye disease (Rieser, Hill, Talor, Bradfield, & Rosen, 1992) or artificial restriction (Mason, 2002) also results in decreased wayfinding performance.

Discussion

Contrary to our prediction, low-vision participants demonstrated more accurate layout knowledge after learning with non-geometric information than with only high-contrast geometric information. Performance after learning in the Photorealistic VE was equivalent to learning with a map or in a real building. Yet, acquisition of layout knowledge required similar amounts of exploration in both types of VEs.

Many participants did not pass the learning criterion with only geometric information. Given more exploration time, a difference might have emerged in the acquisition rate between the two VEs, and the difference in the accuracy of layout knowledge may have disappeared. Either scenario indicates that acquiring layout knowledge is more difficult with only geometric visual information.

Because performance in the Photorealistic VE was comparable to learning with a Map or in the Real building, difficulty acquiring layout knowledge in the Sparse VE cannot be attributed to a lack of familiarity with first-person virtual environments or with the keystroke interface. Therefore, the results must be due to the visual information provided in these displays and how they are used to learn the layout.

Experiment 2 also demonstrated that people with low vision can use visual displays to learn novel layouts; learning with a map or a Photorealistic VE was comparable to learning by walking around in the real space. Both types of computer displays could potentially be used for pre-journey exploration and familiarization with buildings. High-contrast visual maps could be used on-the-fly during travel using a laptop or PDA. Also, low-vision participants performed better with maps than in the Photorealistic VE, a result found previously with normally-sighted participants (Farrell, et al., 2003). This result and the fact that it is easier and less costly to generate maps of buildings suggest that maps are more practical as a visual navigation aid than VEs for the visually impaired.

The broad age distribution of the low-vision participants may be a contributing factor to the reduced performance in the Sparse VE. Our low-vision group included seven participants younger than 50 and six older than 50. All but one of the participants who did not pass the learning test in the Sparse VE were older than 50 years old. This observation suggests that older participants may have had more difficulty acquiring layout knowledge in the Sparse VE. The goal of Experiment 3 was to test the effects of age on our navigation tasks, which helped in interpreting the low-vision results in Experiment 2.

Experiment 3

Visual impairments are more common among the older population, largely because of the prevalence of age-related eye diseases. Non-visual age-related factors might contribute to the ability to use only geometric or additional non-geometric information when learning indoor layouts, as suggested previously by Moffatt & Resnick (2002). Therefore, in Experiment 3 we tested a group of older participants (ages 50-70) with normal vision to explore whether age affects the ability to learn layouts in the Sparse and Photorealistic VEs.

Method

Participants

We tested four males and four females between the ages of 50 and 70 (mean age = 60, SD = 5), all with normal or corrected-to-normal vision. All participants passed the Mini-Mental State Exam (scores ranged from 29-30 out of 30), designed to assess general cognitive functioning. Participants were also asked to rate their health status (1 = poor, 5 = very healthy, median participant rating = 4.5), level of activity (1 = not active, 5 = very active, median rating = 4), driving experience (1 = never drive, 4 = drive daily, median rating = 4), and computer experience (1 = never use computer, 4 = use computer daily, median rating = 3). Participants were not familiar with the psychology building where testing occurred and were monetarily compensated for their time.

Procedure

The materials and procedure were identical to those used in Experiment 1 except that participants were not tested in the Map condition because the primary purpose was to assess how age affects performance in the Sparse and Photorealistic VE conditions.

Results

Acquisition of Layout Knowledge

Rates of acquisition varied significantly by Presentation Mode ($F(2, 14) = 27.973$, $p < 0.001$, $\eta_p^2 = 0.80$). Significantly more exploration was required in the Sparse and Photorealistic VEs compared to the Real building ($p < 0.01$) as shown in Figure 2.11. However, there was no difference in exploration time between the two virtual environments. Furthermore, six participants were not able to pass the learning criterion in the Sparse VE, and four of these participants also did not pass in the Photorealistic VE.

This indicates that some participants found it particularly difficult to learn layouts using desktop virtual environments.

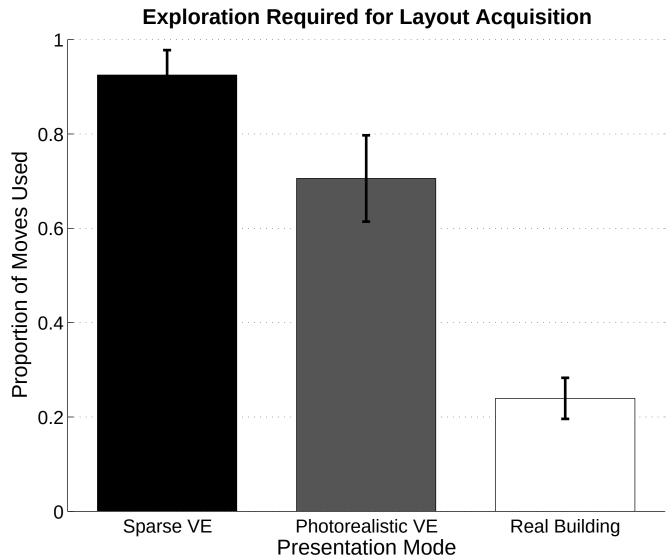

Figure 2.11. Mean proportion of moves used by older normally-sighted participants to reach criterion learning in three presentation modes. Error bars indicate +/- 1 standard error.

Accuracy of Layout Knowledge

Performance on target localization varied significantly with Presentation Mode ($\chi^2(2) = 11.12$, $p = 0.004$, $\eta_p^2 = 0.70$). As depicted in Figure 2.12, performance was significantly worse after learning in the Sparse VE than either the Photorealistic VE ($p = 0.012$) or the Real building ($p = 0.003$). Similar to the low-vision participants, non-geometric visual information was useful for encoding and representing the correct locations of targets in memory. Map drawing performance did not reveal significant differences between presentation modes ($F(2, 14) = 3.263$, $p = 0.069$, $\eta_p^2 = 0.32$), although the trend was still that layout knowledge was least accurate in the Sparse VE (Figure 2.13). There were no significant correlations between individual ratings of computer ability and acquisition rate or accuracy of layout knowledge.

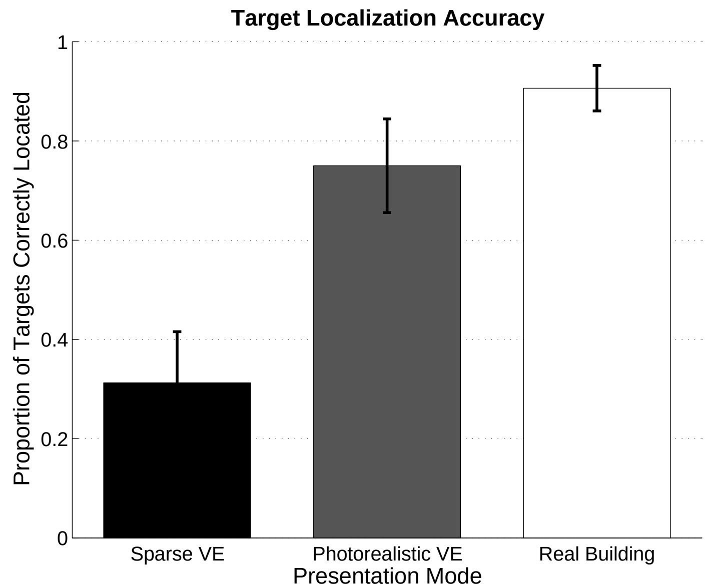

Figure 2.12. Mean proportion of targets correctly localized by older normally-sighted participants in three presentation modes. Error bars indicate +/- 1 standard error.

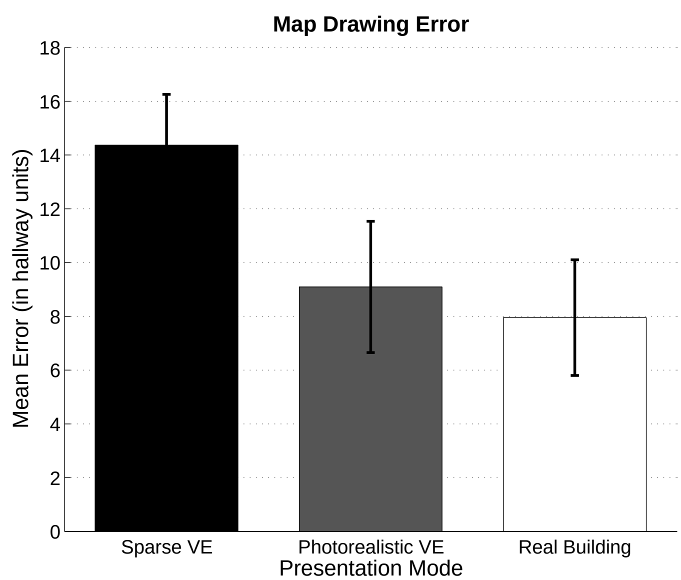

Figure 2.13. Mean error (in hallway units) in the map drawing task by older normally-sighted participants in three presentation modes. Error bars indicate +/- 1 standard error.

Comparing the Effects of Visual Impairment and Age on Performance

The results of the older normally-sighted adults showed a similar pattern to the low-vision group (Experiment 2). The acquisition rates for both groups showed that the Sparse VE required the most exploration to pass the learning criterion, but not significantly more than the Photorealistic VE. Furthermore, both the low-vision and older groups demonstrated significant differences in overall layout knowledge between presentation modes. Target localization accuracy was most affected by the visual information available during learning; both groups performed significantly worse after learning in the Sparse VE compared to the Photorealistic VE.

Figure 2.14 combines target localization data from all three experiments in a plot of performance by visual ability and age. It indicates that individuals in the same age group had similar target localization scores regardless of visual ability. We performed Wilcoxon Rank Sum tests to evaluate the effects of Visual Condition and Age, and a Friedman test to evaluate the effect of Presentation Mode. The results revealed a significant effect of Age ($p = 0.011$) but not Visual Condition. There was also a highly significant effect of Presentation Mode ($\chi^2(2) = 26.14$, $p < 0.001$) with significant differences between the Sparse VE and the other conditions (Bonferronized pairwise comparisons: $p < 0.001$). Wilcoxon Rank Sum tests between the performance of younger and older individuals in the three Presentation Modes only revealed a significant effect in the Sparse VE ($p = 0.002$) but not in the other conditions. Together, these results suggest that age, rather than the conjunction of age and vision loss, drives performance in the Sparse and Photorealistic VEs. Older people find it much more difficult to learn layouts based solely on geometrical information, and therefore seem to benefit more from the addition of non-geometrical cues.

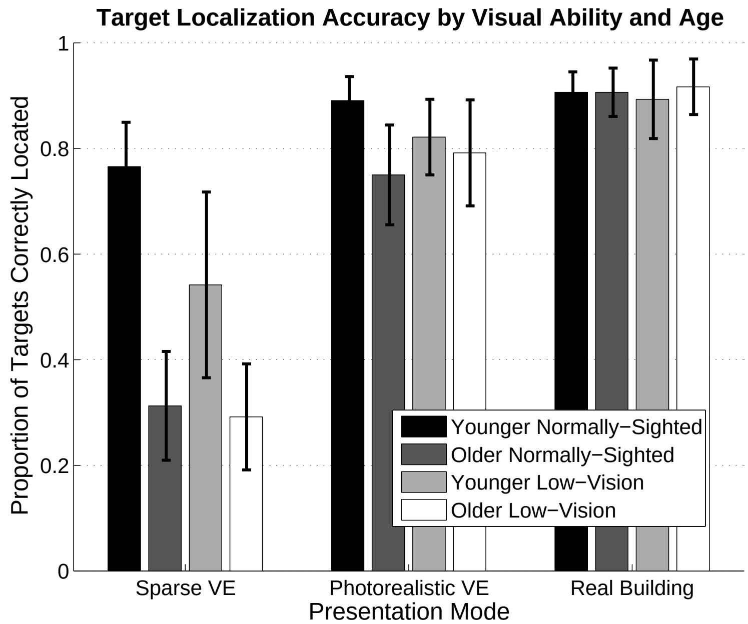

Figure 2.14. Mean proportion of targets correctly localized (+/- 1 SE) for younger and older adults with normal and low vision.

<u>Discussion</u>

These results indicate that age does influence the ability to learn layouts in virtual environments. Older participants had difficulty on tasks assessing layout knowledge, especially target localization, after learning in the Sparse VE, which is consistent with previous studies (Moffatt & Resnick, 2002). This indicates that the decreased performance of low-vision participants in the geometrical environments (Experiment 2) could be at least partially due to the inclusion of older participants.

Aging is related to declines in declarative learning, which is influenced by working memory capability (Kirasic, Allen, Dobson, & Binder, 1996). The demands of trying to disambiguate position with only geometric visual information may impose more cognitive load than can be handled by older people, resulting in a decreased ability to encode layout information. Previous research has demonstrated that even younger adults exhibit less than optimal navigation performance compared to an ideal observer because of limitations in remembering their path in geometric environments (Stankiewicz, Legge, Mansfield, & Schlicht, 2006). Older individuals may be even more affected by memory limitations while navigating and may rely more on non-geometric cues that require fewer cognitive resources when encoding locations in a mental representation of a layout. This coincides with the results of previous experiments that have suggested that older adults reproduce target locations more accurately with object landmark information than with only geometric information (Moffatt & Resnick, 2002).

Non-visual information, associated with walking in the real building, was advantageous for older normally-sighted adults, allowing them to learn layouts more quickly. These cues seem to facilitate memory for spatial layout, perhaps by reducing the cognitive load required to integrate information over multiple views in the vision-only environments.

General Discussion

These experiments investigated the types of visual information needed by younger and older normal and low-vision individuals when learning novel indoor layouts. By comparing performance in Sparse and Photorealistic VEs, we specifically tested whether geometric information by itself (Sparse VEs) was sufficient for learning layouts from a first-person perspective or whether the rendering of additional non-geometric information (Photorealistic VEs) was advantageous. We measured both the acquisition and accuracy of layout knowledge after learning in the different presentation modes (Sparse VE, Photorealistic VE, Map, and Real building).

Resolving Ambiguous Geometric Visual Information

All groups of participants found it more difficult to learn layouts with only geometric information; the addition of non-geometric information improved acquisition and/or accuracy of layout knowledge. Geometric visual renderings can result in ambiguity about locations in a layout. For example, the rectangular environments described earlier had identical geometric structure at opposite corners that were easily confusable for rats (Cheng, 1986) or children (Hermer & Spelke, 1996). The same geometrical ambiguities can exist in more complex environments. For example, suppose a layout, rendered only with geometrical hallway structure, has two T-junctions. When you arrive at a T-junction, how do you know which one it is?

Consider two options for resolving the ambiguity: 1) Perceptual solution: The two T-intersections may differ in the distances to adjacent intersections and the branching patterns at the adjacent intersections. If you encoded geometric information to this level of detail during layout acquisition, you can use it to resolve the ambiguity; 2) Path memory: You can resolve the ambiguity if you remember your previous location and the path taken to reach the current T-junction, assuming a unique path is required to reach each T-junction from the previous location. Both of these techniques provide means for spatial updating within purely geometrical layouts, but require demanding perceptual encoding and/or memory processes.

It is plausible that impaired vision or natural aging could reduce the capacity to accomplish perceptually or cognitively demanding spatial updating. Individuals with visual impairments may not be able to detect perceptual differences in layout structure, and older adults may have reduced memory capacity. If so, a third kind of mechanism, the presence of additional redundant non-geometrical cues could be helpful. For instance, one of the two T-intersections might have a water fountain that could resolve the geometrical ambiguity if either of the two strategies just outlined could not be used. Thus, non-geometric information may be especially useful for people with visual impairment or who are older.

Conclusions

The acquisition of layout knowledge was clearly influenced by the visual information available during exploration for all groups of participants. Non-geometric visual cues reduced the amount of exploration needed to learn a novel layout, especially for younger normally-sighted individuals.

Non-geometric information improved the accuracy of layout knowledge for both the low-vision and older normally-sighted groups, but not for younger normally-sighted individuals. Low-vision and older sighted people especially benefited when localizing targets in the real building after learning layouts with non-geometric information.

It is likely that age was a contributing factor in explaining the relatively poor performance of the low-vision participants when only geometric information was available during learning, as indicated by Figure 2.14. One implication is that degradation in visual function does not necessarily lead to reduced performance. Also, considering that the low-vision population is older, this age effect is important in determining the types of visual displays that will be useful navigation aids for people with low vision.

We also compared learning in the first-person VEs to learning with Maps and in a Real building. Learning in the Map condition resulted in similar performance compared to the Photorealistic VE and the Real building in all three experiments, suggesting no advantage to having global layout information when visually-rich information from a first-person perspective is available. Non-visual information was only advantageous for older normally-sighted adults.

Finally, we demonstrated that people with low vision can learn layouts as effectively using Maps and Photorealistic VEs displayed on a desktop computer as walking through a real building. Yet, older individuals may not learn as quickly in VEs as in real buildings.

In conclusion, this study showed that information about layout geometry is sufficient for learning complex indoor layouts, at least by younger individuals with normal vision. Additional non-geometric information aids learning by low-vision and older people. The reason for this reliance on non-geometric information may be due to the increased cognitive demands required to create and use cognitive maps with only geometrical information. A positive practical outcome was that individuals with low vision can use virtual displays to learn layouts prior to going to a real building, which may have applications for developing indoor navigation aids.

Chapter 3:

Assessment of Indoor Route-finding Technology for Visually–Impaired People

Introduction

A common problem faced by individuals with visual impairment is navigating through unfamiliar buildings. Although canes and guide dogs aid in obstacle avoidance, it is still difficult to plan and follow routes to specific locations. Our laboratory is developing an indoor wayfinding technology, called the Building Navigator, that contains a route-planning feature. This chapter describes user tests of this feature.

Successful wayfinding requires two main pieces of information: 1) the current location and heading of the individual, and 2) the route to the destination. Routes consist of waypoints, locations where the navigator changes direction. To follow a route, the navigator needs to have real-time access to information about the distance and direction to waypoints until the destination is reached.

Thus far, development of indoor wayfinding technology for visually-impaired travelers has addressed the first issue of localization. Braille signs are now commonly used in buildings, but they are sometimes difficult to locate because of sporadic placement. Also, many people with visual impairment do not read Braille. Various other positioning technologies have been explored, such as Talking Signs (www.talkingsigns.com), Talking Lights (www.talking-lights.com), RFID tags, and systems using wireless signals, but there are still limitations in the accuracy and the cost of installing and maintaining these systems (Giudice & Legge, 2008; Loomis, Golledge & Klatzky, 2001). Our Building Navigator software has been designed to be integrated with positioning technologies while also providing information about the layout and other salient features of indoor spaces.

The Building Navigator

The Building Navigator provides information about the spatial layout of rooms, hallways, and other important features in buildings through synthetic speech output. This software was designed to be part of a portable system, perhaps installed on a cell phone or PDA. The following paragraphs describe the major components of the system: the Building Database that contains information about the layout of the floor (digital map), the Floor Builder used to input building layout information into the database, and two interface components for exploration and route finding.

The Building Database

The Building Database stores information about the physical features of the building as well as the spatial layout of these features. Stored features include spaces, such as rooms and lobbies, as well as important objects such as water fountains. Features are encoded into the database by first dividing a floor layout into meaningful spatial regions (Figure 3.1). These regions are assigned to feature types (e.g., door, room, hallway, window, stairs, elevator), and feature types are grouped into broader categories (e.g. physical space, connecting space, utility features) to facilitate fast searches for layout information.

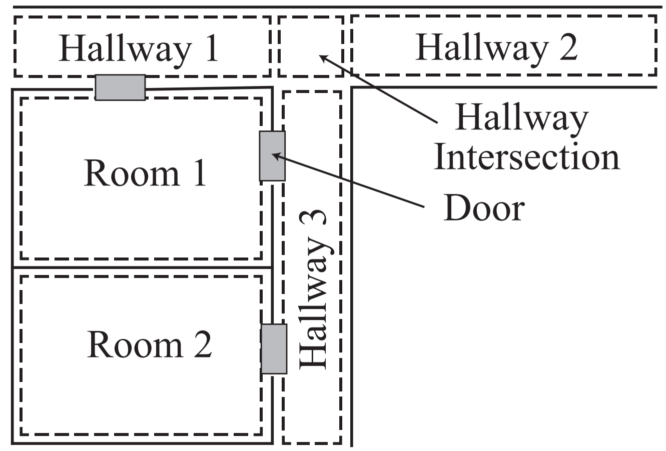

Figure 3.1. An example of how a floor is segmented into features when stored in the Building Database.

Features that are spatially adjacent in the layout are associated with each other in the database through a set of logical relationships. For example, to indicate that a room is accessible from a particular hallway, a door feature is associated with both the hallway and room features. This makes the door logically accessible from both the hallway and the room.

The Building Database also includes functions for acquiring information (e.g. getting a list of known buildings, floors within a building, types of features present within a building) and also for managing information requests from input and output plugins. Input plugins are intended to handle environmental sensors like a wireless network location device, dead reckoning systems, or rotation/orientation sensors. Their primary purpose is to gather information regarding the user’s location, heading and movements. Output plugins provide an interface through which the user interacts with the Building Database and the rest of the navigation device. This paper discusses two types of speech enabled output plugins for use in exploration and route finding.

Entering Building Information into the Database

A separate software program, called Floor Builder, is used for data entry. Conventional spatial mapping applications do not encode the range of features used by the Building Database, nor their spatial relations, resulting in the need for a custom application. First, a map of the building is digitized and segmented into features by a human operator, currently an experimenter. In the future, this person would be someone trained in the use of the software, and not the visually-impaired end user. The Floor Builder software has a graphical-user interface that allows the operator to follow along a series of simple point-and-click steps to parse the map into features. The segmented map is converted into the necessary data structures and uploaded into the Building Database. With the current software, an experienced operator can complete the mapping of a floor with 50 rooms in about two hours.

User Interfaces for Navigation

The Building Navigator presents information via synthetic speech output, but in principle the same verbal information could be sent to a Braille display. A significant challenge was developing verbal descriptions of the space that were concise, informative, and easy to understand. These descriptions benefited from prior work on verbal descriptions for outdoor wayfinding, but differ in important ways. For instance, unlike streets, indoor hallways are not typically named. Also, more information needs to be conveyed within a smaller area of space in indoor environments compared to outdoor environments. Previous studies have found that blindfolded sighted and blind individuals are able to effectively learn and navigate through buildings using consistent and structured verbal descriptions of layout geometry (Giudice, 2004; Giudice, Bakdash &

Legge, 2007).

Upon entering an unfamiliar building, a user may have two possible goals: 1) to become familiar with the overall building layout through exploration, or 2) to find a specific location by following a route. The Building Navigator currently supports these two types of navigation.

Exploration Mode

In the Exploration Mode, speech output describes the layout of features near the user's current location. For example, if the user is standing in a lobby, the set of features described in this mode would include the doors and hallways located on the perimeter of the lobby. Users can also receive egocentric and allocentric descriptions of these features. An egocentric description provides the direction to a feature relative to the user's current location and chosen heading. Allocentric descriptions present information with respect to a set of absolute reference directions such as North, South, East and West. These descriptions also provide distances to features. See Figure 3.2A for an example of a feature list and the corresponding egocentric and allocentric descriptions.

The Exploration Mode can be used for virtual exploration of a space, meaning a user can simulate navigation through the layout without setting foot in the physical building. The interaction is similar to having a guide provide detailed layout information at the user's request. This ability to virtually explore and learn a space before actually going there, called pre-journey learning, has proven beneficial to blind and visually-impaired navigators (Holmes, Jansson & Jansson, 1996).

Route Mode

The Route Mode produces a list of instructions for navigating from a start to goal location via a series of waypoints (Figure 3.2B). The route is computed using Dijkstra's Shortest Path algorithm (Dijkstra, 1959), a well-known graph search algorithm that finds the shortest path between nodes. The user listens to one route instruction at a time, and selects the instruction by moving up or down in the list with key presses. The first instruction always indicates the user's starting location and heading. The subsequent instructions describe the distance and direction of travel to a series of waypoints, and the direction of the turns. The final instruction indicates the distance to the goal location and on which side of the space it is located (for example, south side of the hallway). Currently, all directions use an allocentric frame of reference (i.e. North, South, East, and West). To provide egocentric directions, the system must incorporate sensors or otherwise obtain information about the user's current heading.

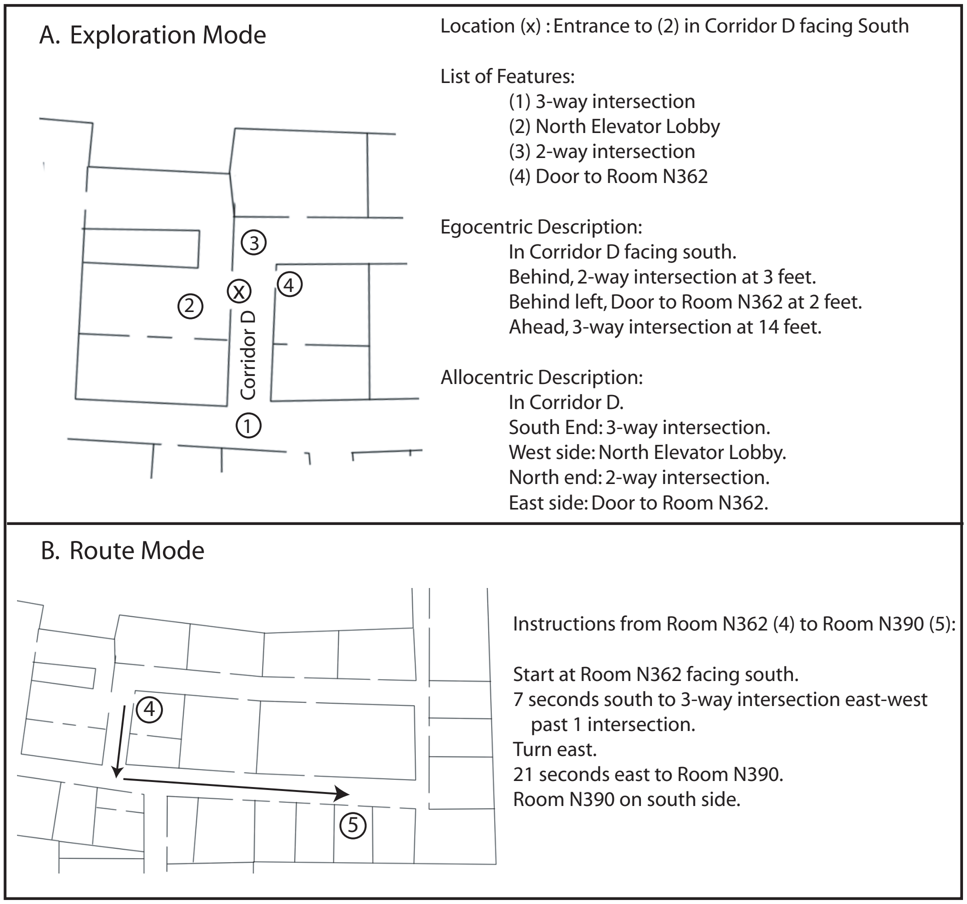

Figure 3.2. A) Example of a feature list, egocentric, and allocentric descriptions provided by the Exploration Mode of the Building Navigator. B) Example of route instructions provided by the Route Mode of the Building Navigator.

A sample verbal description to a waypoint is “104 feet south to 3-way intersection west past 2 intersections.” The descriptions contain three critical parts in a standard format. First, the distance and direction to the waypoint is given (“104 feet south”). Distance is provided in one of three units: feet, the number of steps, or the time in seconds to travel. Distances in number of steps and seconds are individually calibrated according to step length and walking speed.

The second part of the instruction describes the geometry of the waypoint intersection (“3-way intersection west”). Figure 3.3 depicts examples of possible intersections and how they are described by the Building Navigator. The waypoint can be a 2, 3, or 4-way intersection. Also, for 2 and 3-way intersections, the directions of the branching arms are described in an allocentric frame of reference. Consequently, the user is responsible for translating the description into an egocentric frame of reference (e.g. 3-way intersection west means there is a hallway branching to the left if the user is facing north, or a hallway branching to the right if the user is facing south). Lastly, the verbal description indicates how many intersections the user will pass before coming upon the waypoint ("past 2 intersections"). While these descriptions initially seem hard to parse, their structured format supports rapid learning. Participants learned to comprehend the descriptions with modest amounts of practice.

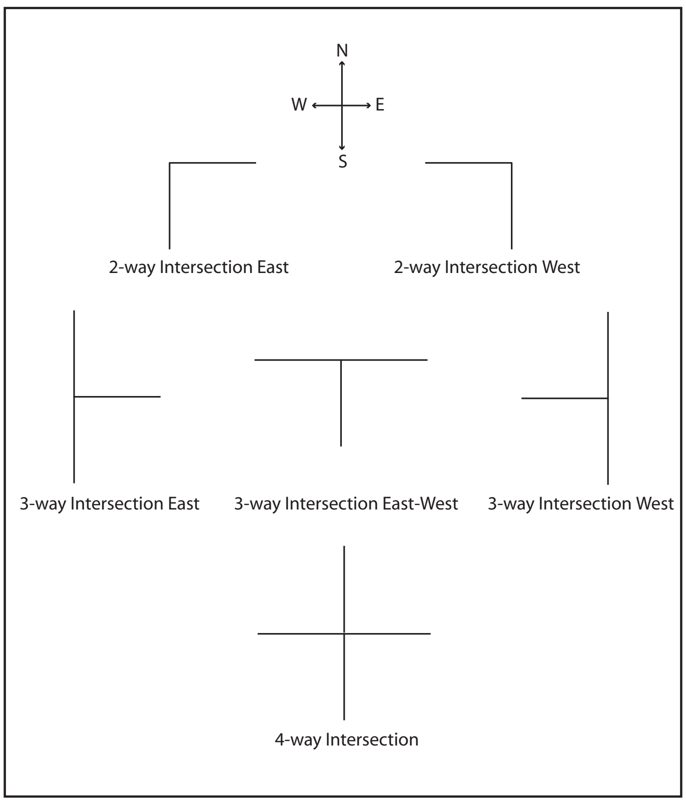

Figure 3.3. Sample descriptions of 2-way (L-junction), 3-way (T-junction), and 4-way (+ junction) intersections provided in the Route Mode. Each branching arm represents a hallway.

The remainder of this paper describes user testing of the Route Mode of the Building Navigator. The goal was to validate the system as a useful application for improving wayfinding by visually- impaired people.

User Testing of the Route Mode of the Building Navigator

We tested the ability of blindfolded normally-sighted and visually-impaired individuals to use the Route Mode of the Building Navigator interface without the use of positioning sensors. The goal was to determine if the route information provided by the Building Navigator resulted in improved wayfinding performance. A second goal was to compare route-following performance with three metrics for describing distances to waypoints (feet, steps, or seconds).

Method

Participants

Twelve normally-sighted (6 males and 6 females, age range of 19-29) and eleven visually-impaired participants were tested. The criteria used for selecting visually-impaired participants were that their impairment resulted in limited or no access to visual signs, they were no older than 60 years of age, and otherwise had no deficits that impaired mobility. Table 3.1 describes additional characteristics of the visually-impaired participants. All participants provided informed consent and were compensated either monetarily or with extra credit in their introductory psychology course. This study was approved by the Institutional Review Board at the University of Minnesota.

Apparatus and Materials

The Building Navigator software was installed on an Acer TravelMate 3000 laptop carried by participants in a backpack. Participants wore headphones connected to the laptop to hear the speech output, and used a wireless numeric keypad to communicate with the laptop. The experimenter could also communicate with the user's laptop (via a Bluetooth connection to a second laptop) for the purpose of entering start and goal locations for wayfinding trials. Normally-sighted users wore a blindfold during testing.

	Gender	Age	Age of onset	Low vision or blind	Mobility	logMAR acuity	Diagnosis
1	F	35	not available	Blind	Dog	NA	not available
2	F	35	Birth	Blind	Dog	NA	Clouded Cornea, Microthalmas
3	M	34	6 months	Blind	Dog	NA	Retrolental Fibroplasia
4	M	43	19 years	LV	Cane	1.77	Retinitis Pigmentosa
5	F	26	Birth	LV	Cane	1.32	Advanced Retinitis Pigmentosa
6	F	41	18 months	Blind	Dog	NA	Retinal Blastoma
7	M	24	Birth	Blind	Cane	NA	Retinopathy of Prematurity
8	M	47	Birth	LV	none	1.18	Congenital Cataracts
9	M	60	6 years	LV	none	1.7	Secondary Corneal Opacification
10	M	52	Birth	LV	none	1.44	Glaucoma, Congenital Cataracts
11	F	55	Birth	Blind	Dog	NA	Retinopathy of Prematurity

Table 3.1. Description of the eleven visually-impaired participants.

Procedure

Participants were tested in four conditions using a within-subjects design. In three conditions participants used the technology, once in each distance mode (feet, steps or seconds). In a baseline condition, participants were not allowed access to the Building Navigator technology.

In all conditions, participants were allowed to ask “bystanders,” played by the experimenter, for information. Bystander queries could only be made at office doors. The “bystander” provided participants with their current location and the egocentric direction to travel to reach the destination (e.g., “You are at Room 426. Go right.”). Participants were instructed to minimize the number of bystander queries and to only make them when necessary, as if they were interrupting people in their offices. In the real world, when signage is not accessible, visually impaired individuals have no recourse but to seek information from bystanders. We simulated bystanders, rather than relying on actual bystanders, to equalize the access to information across participants and conditions.

Each condition was tested in a different building layout, and the condition-layout pairings and the order of the conditions were counterbalanced. Participants were first trained to use the system, as described below, before proceeding with testing.

Calibration

To individually calibrate distances given in steps and seconds, we obtained an accurate estimate of each participant’s step length and walking speed. Participants were asked to wear the backpack with the laptop inside, as during testing. Sighted participants practiced walking blindfolded while guided by an experimenter until they felt comfortable. All participants were asked to walk a thirty-foot length of hallway three times while the experimenter counted their steps and timed them. The average number of steps and time walked were then entered into the system. Normally-sighted participants had a mean step length of 1.85 feet/step and a mean velocity of 2.92 feet/second. Visually-impaired participants had a mean step length of 1.98 feet/step and a mean velocity of 3.43 feet/second. T-tests indicated no significant difference between groups for step length, but a marginally significant difference in velocity ($p = 0.02$). The use of step length and walking speed to compute travel distances relies on consistency in a subject's walking characteristics. Previous work in our lab has shown that step-length variability is small for both blind and sighted individuals (Mason, Legge, & Kallie, 2005).

Training

Participants were first introduced to the structure of route descriptions produced by the system. This included explicit training on verbal descriptions used to convey the geometry of intersections (refer to the examples in Figure 3.3). The experimenter showed an example of each type of intersection, using tactile maps for visually-impaired participants, and explained the corresponding verbal description given by the system. The experimenter then tested participants on their understanding by asking them to describe the geometry of intersections corresponding to verbal descriptions.

Participants were also trained on the functions of the keypad. The "2" and "8" keys were used to move up and down in the list of instructions. When in the "seconds" distance mode, the "4" was used to start and pause a timer that beeped until the number of seconds indicated by the selected instruction elapsed. The "6" was used to stop and reset the timer. Lastly, the "/" was used to repeat an instruction. Participants practiced using the keys with a sample list of instructions, and thus were also familiarized with the speech output.

To further familiarize the participant with the task and technology, participants completed three practice routes, one in each distance mode. The practice routes were located on a floor other than those used for testing.

Testing

Participants were tested on four routes in a novel layout for each condition. The routes for each layout were chosen to be of similar difficulty. The average distance per route was 144.8 feet (*SD* = 8.3) and the average number of turns required was 2 (*SD* = 0.24). Figure 3.4 shows a sample set of four routes used for testing.

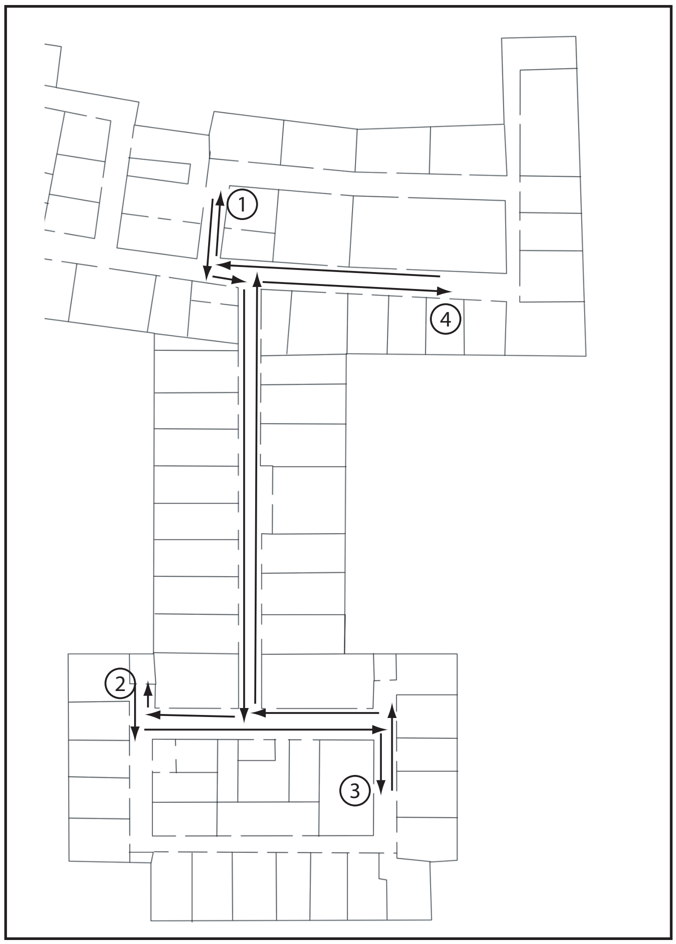

Figure 3.4. Example of a layout with four routes used for testing.

At the beginning of a trial, participants were led to the starting location and instructed to face a specific direction. The experimenter then stated their current location, the direction they were facing, and the goal location (e.g. "You are at room N362 facing south. Go to room N349."). Participants then attempted to find the goal location, using bystander queries when necessary, and indicated when they thought they had arrived at the goal location. The trial ended when participants correctly found the goal location or when they gave up. Visually-impaired participants also ended trials in the technology conditions when they felt the system was no longer helpful and they chose to rely exclusively on bystander queries. This was to prevent these participants, who were on average older than the normally-sighted participants, from getting too frustrated or tired during the experiment. The beginning of the next trial started at the previous goal location.

The experimenter told blindfolded sighted participants when they passed intersections, but visually-impaired participants were not provided with this information to better simulate a real world context. Also, if participants deviated from the prescribed route, they did not receive a new set of instructions from the Building Navigator describing the route from their current location to the goal location. A second experimenter timed each trial with a stopwatch, recorded the participant’s trajectory through the layout, and noted the location of bystander queries. At the end of the experiment, participants completed a survey asking them to evaluate the technology and to rank the conditions from most to least preferred.

Data Analysis

Wayfinding performance was evaluated using several measures. For each condition, we computed the average number of turns made, distance traveled, and time taken to complete a route. We also measured the average number of bystander queries made in each condition as an indicator of how independently participants could locate the rooms.

The results for the normal and visually-impaired participants were analyzed separately. The dependent measures--number of turns, distance traveled, and time--were analyzed using Analyses of Variance (ANOVA) blocked on subject. Because the data were not normally-distributed, Box-Cox power transformations were performed on the data. The nonparametric version of the ANOVA was conducted for the bystander query measure since the data could not be normalized with a transformation. For all measures, contrasts comparing conditions with the technology to the baseline condition were performed. Also, Bonferroni-corrected pairwise comparisons were conducted to evaluate if one of the three distance modes provided by the technology resulted in better performance.

Results

Participants with Normal Vision

Participants were able to find 100% of the rooms across conditions. For all measures (Figure 3.5), the ANOVAs were highly significant ($p < 0.004$) and performance was significantly better with the technology than in the baseline condition ($p < 0.007$). For three of the four measures (number of turns made, distance traveled, and number of bystander queries), performance in each distance mode was significantly better than the baseline condition ($p < 0.001$). When distance was given in feet or steps, participants took significantly less time finding rooms compared to baseline ($p < 0.003$). There were no significant differences between distance modes for any of the measures. As shown by the median rankings displayed in Table 3.2, there was no clear preference for a particular distance mode when using the technology.

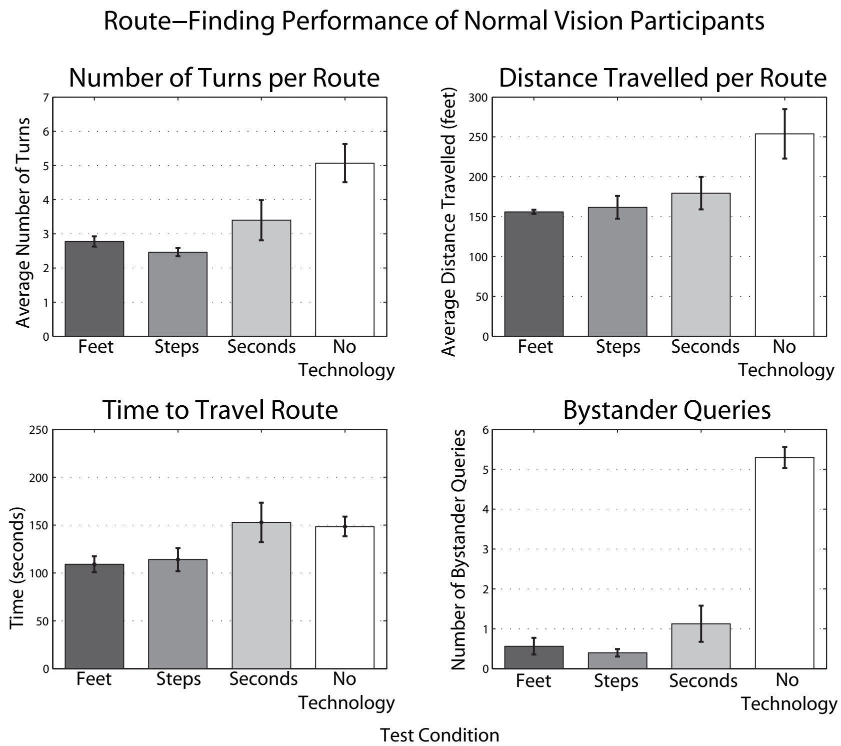

Figure 3.5. Performance of normal vision participants in the route-finding task as measured by the number of turns made, distance travelled, time travelled, and number of bystander queries made.

Visually-Impaired Participants

Averaged across conditions, participants were able to find 93% of the target rooms. ANOVAs revealed significant effects for the number of turns made, distance traveled, and number of bystander queries ($p < 0.04$) (Figure 3.6). For these measures, performance was significantly better with the technology compared to the baseline condition ($p < 0.02$).

Participants made significantly fewer bystander queries in all the distance modes compared to the baseline condition ($p < 0.001$). They made significantly fewer turns with distance in steps and seconds ($p < 0.008$). When distance was given in steps, they also traveled a significantly shorter distance compared to baseline ($p = 0.006$). There were no significant differences between distance modes in any of the measures. Table 3.2 indicates that participants preferred distance in steps when using the technology.

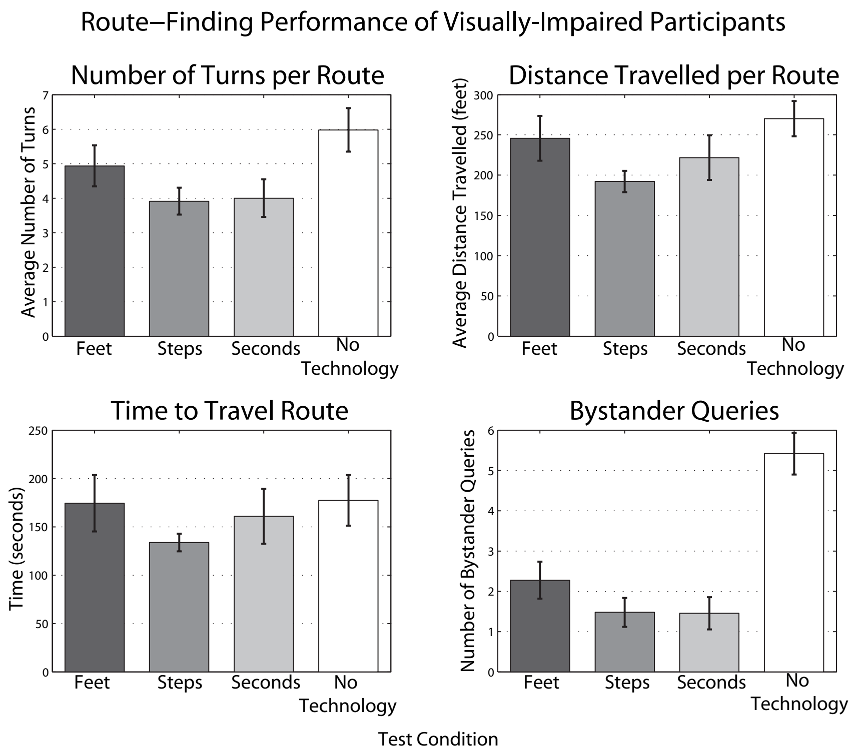

Figure 3.6. Performance of visually-impaired participants in the route-finding task as measured by the number of turns made, distance travelled, time travelled, and number of bystander queries made.

	Rankings for each condition (1 = most preferred)			
Vision Group	**Building Navigator: Distance in feet**	**Building Navigator: Distance in steps**	**Building Navigator: Distance in seconds**	**No technology**
Normal	2.5	2	3	4
Impaired	3	1	3	3

Table 3.2. Median participant preference rankings of the four test conditions

Discussion

The goals of this experiment were to test if 1) the route-finding technology improved the ability to find rooms in an unfamiliar building, and 2) if one of the three distance modes (feet, steps or seconds) provided by the technology resulted in better performance. Four measures were used to evaluate route-finding performance: the number of turns made, the distance travelled, the time taken to find a room, and the number of bystander queries.

For normal vision participants, use of the technology improved performance on all measures. For visually-impaired participants, the technology improved the ability to take the shortest route to the target room, as indicated by a reduction in the number of turns and distance travelled. The technology also allowed these participants to navigate more independently, demonstrated by fewer bystander queries. The technology did not significantly decrease the time needed to complete the routes, perhaps because of the extra time required to interact with the technology (e.g., scrolling through the list of waypoints and absorbing the verbal instructions.).

The rankings provided by normally-sighted participants in the post-experiment survey did not reflect a strong preference among the distance modes. Visually-impaired participants preferred distance in steps compared to the other distance modes. These participants also performed better in the route-finding task with distance in steps, demonstrated by fewer turns and shorter travelled distances. According to comments from the participants and observations during testing, some visually-impaired participants did not have a good understanding of distance in feet. Distance in seconds was not always reliable because walking speed was variable, especially for participants with guide dogs. Most participants preferred distance in steps because it was consistently accurate, likely due to low variability of step length (Mason, et al., 2005), and they had more control over the counting compared to when distance was given in seconds.

Some of the visually-impaired participants thought it was useful to know how many intersections to pass before arriving at a waypoint while others did not seem to use the information. The usefulness of this information seemed to depend on whether the user could detect intersections, either using their residual vision or with their cane or guide dog. Several participants commented that it was difficult to maintain orientation in allocentric coordinates (North, South, East, West), and that they would have preferred egocentric directions. Indeed, most mistakes participants made were turning the wrong direction when following the route instructions. These orientation issues will be solved in the future when the Route Mode is integrated with positioning and heading sensors.

In this study, the baseline condition required wayfinding without the Building Navigator technology. This baseline provided participants with access to bystander information at every doorway in the layout. We expect the assistive technology to be even more advantageous in realistic situations when access to bystander information is less frequent.

We conclude that route instructions can improve wayfinding by individuals with visual impairment in unfamiliar buildings. Even without additional positioning sensors, visually-impaired participants were able to successfully follow instructions to locate rooms. Additional findings indicate that distances to waypoints can be conveyed effectively by converting metric distance into an estimate of the number of steps by the user. The preference for step counting to estimate distance and the improved performance with this metric is an important finding for the design of wayfinding technology for visually-impaired people.

Chapter 4:

Integration of Visual and Walking Information for Localization Depends on the Congruence and Reliability of Cues

Introduction

Humans, and animals, can use two possible sources of information to navigate through the world: landmarks and path integration (Gallistel, 1990; Loomis, et al., 1999). Together, these mechanisms allow humans to build and reference mental representations of environments for localization. The current study explores how two factors, information congruence and reliability, influence the interaction of visual landmarks and path integration. We test two predictions: 1) visual landmark and path integration information are combined to determine position only when they provide compatible information, and 2) when combined, dependence on path integration increases as the reliability of visual landmark information decreases. The reliability of visual information is manipulated through blur.

Visual Landmark Guidance versus Path Integration

Path integration refers to methods of position-updating by integrating one's velocity and acceleration over time (Loomis, et al., 1999). In humans, angular and linear displacements are measured using external signals provided by the environment, such as optic flow, and internal signals such as kinesthetic and vestibular signals. Optic flow and kinesthetic cues can be used to determine the velocity of travel, whereas vestibular cues indicate acceleration of the head during movement. Even cognitive strategies in humans, such as step-counting, are adjusted online using path integration mechanisms (Kalia, Schrater, Legge & Kallie, 2008).

The computational mechanism underlying path integration is a vector summation of changes in direction and distance over a series of movements. Each measurement of angular and linear displacement is imperfect because of noise in the sensory signals.

Errors accumulate and result in increasing position uncertainty over time (Etienne, Maurer, & Seguinot, 1996; Gallistel, 1990). This uncertainty can be reduced by referencing a pre-existing cognitive map of the environment consisting of landmarks (Etienne, et al., 2004). The cognitive map associates landmarks with specific positions in the environment.

Humans predominantly use vision to detect landmarks for navigation. Landmarks are typically defined as objects in the environment that are salient, stable, and informative about location (Stankiewicz & Kalia, 2007). Positions of visual landmarks are determined using numerous depth cues, including ocular cues such as accommodation and convergence, and properties that arise from the retinal images such as binocular disparity, familiar size, relative size, and texture gradients. Head motion allows for motion parallax cues to depth while the observer is stationary. Although most depth cues only provide information about relative depth (Palmer, 1999), absolute distances can be computed from the familiar size of objects (Palmer, 1999) and from the angle of gaze to the ground plane if the observer's height is known (Gibson, 1950; Ooi, Wu & He, 2001). By viewing the environment, humans can build a mental representation consisting of the locations of landmarks. This "cognitive map" can then be referenced as the observer moves through the environment. Yet, the accuracy and precision with which landmarks are encoded depends on the quality of visual information. One goal of this study is to explore how degraded visual information affects the precision with which landmarks are encoded.

Integrating Visual Landmark and Path Integration Information

Under what circumstances do visual landmarks reduce the positional uncertainty of path integration? One potential factor is the congruency of the positioning information provided by landmarks and path integration. For example, in an experiment in which landmark and path integration provided conflicting information about the location of the nest, hamsters relied on the landmark for small conflicts and path integration for large conflicts (Etienne, Teroni, Hurni, & Portenier, 1990). When the discrepancy between landmark and path integration was large enough, and therefore incongruent, the hamsters no longer used the landmarks and reverted to path integration. The left side of Figure 4.1 defines congruency as the extent of separation between the distributions of positions indicated by landmarks versus path integration. When these distributions sufficiently overlap, landmarks and path integration provide congruent information. In the current study, we predict that visual landmark and path integration information are integrated only when they are perceived to be congruent (Figure 4.2).

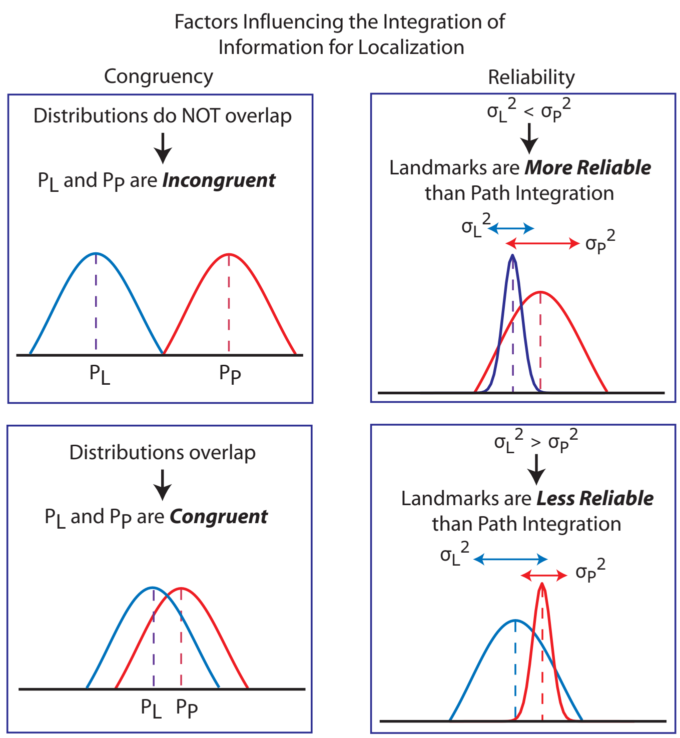

Figure 4.1. Illustration of two factors that may influence the integration of visual landmark and path integration information. These factors depend on the positions indicated by landmarks (P_L) and path integration (P_P), and the uncertainty associated with each strategy (σ_L^2 and σ_P^2).

A second potential factor influencing the integration of landmarks and path integration is their reliability, the inverse of the variability of the estimates with each cue (right panel of Figure 4.1). Maximum Likelihood Estimation (MLE) formalizes how reliability may influence the estimation of position when two sources of information are available. According to this statistical model, when a property in the world, *S*, is estimated from observations of multiple types, for example a visual (*V*) and walking (*W*) cue, the probability of *S* given the likelihood of the observed data can be expressed as:

$$P(S \mid V, W) \propto P(V \mid S) \times P(W \mid S) \qquad \text{Equation 4.1}$$

Assuming that the estimates from individual cues are independent and normally-distributed, the maximum likelihood estimate, $\hat{s}$, is equivalent to the sum of the individual estimates weighted by the inverse of their variances:

$$\hat{s} = w_v \hat{s}_v + w_w \hat{s}_w \qquad \text{Equation 4.2}$$

$$w_j = \frac{1/\sigma_j^2}{1/\sigma_v^2 + 1/\sigma_w^2} \qquad \text{Equation 4.3}$$

The weighting indicates the reliability of the cue estimate. Accordingly, the final estimate is closer to the more reliable cue, and its variance is smaller than the variance of the individual estimates.

Several studies in visual psychophysics have determined that human behavior follows the predictions made by cue combination within and across modalities (Ernst & Bülthoff, 2004; Kersten, Mamassian, & Yuille, 2004; Landy, Maloney, Johnston, & Young, 1995; Mamassian, Landy, & Maloney, 2002; van Beers, Sittig, & Gon, 1999). The current study proposes that observers combine visual landmark and path integration information, when they are congruent, according to their reliability to arrive at a statistically optimal estimate of position (Figure 4.2). Based on this model, we predict that path integration is relied upon more for position estimation when visual landmark information becomes unreliable.

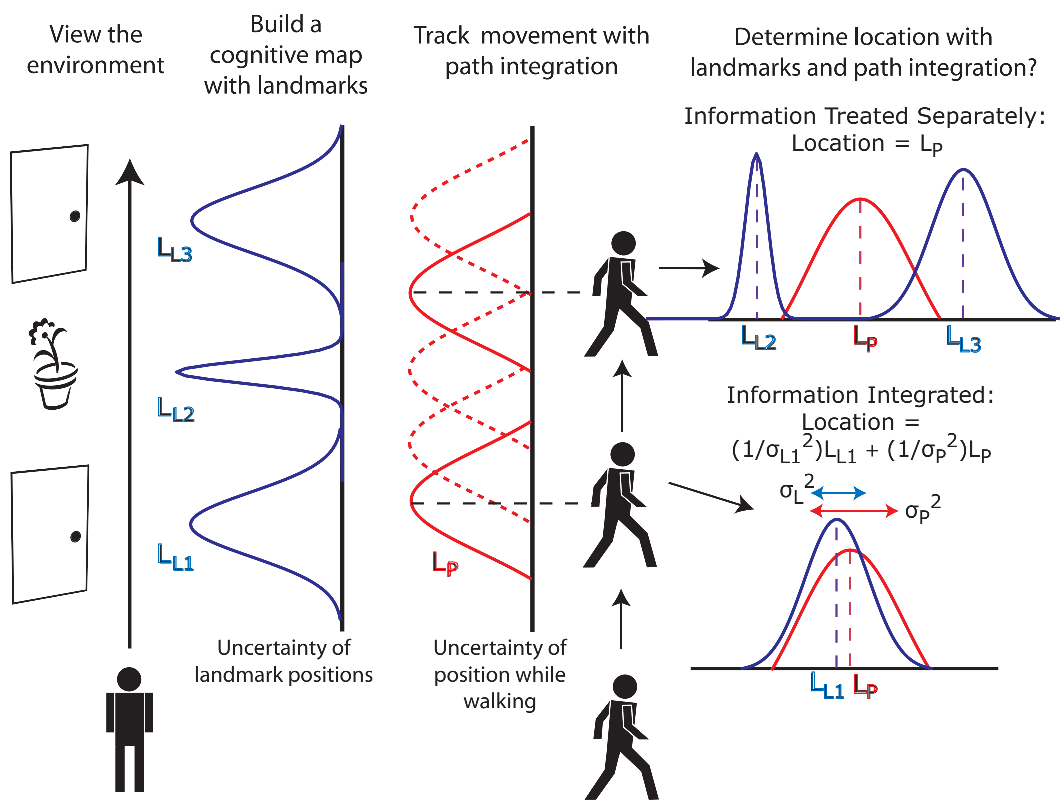

Figure 4.2. Illustration of when and how humans may integrate visual landmark and path integration to determine their location. While viewing a hallway, the observer obtains a cognitive map of the environment including the positions of landmarks. When walking through the environment blindfolded, the observer uses path integration to determine current position with some uncertainty. We predict that in the current study observers will estimate their position based on path integration when this information is incongruent with landmarks. When landmark and path integration information are congruent, observers will integrate these cues according to their reliabilities.

Combining Cues for Localization

Recently, several studies have investigated how humans combine information from different sensory modalities to make spatial judgments in large-scale environments (Ellard & Shaughnessy, 2003; Harris, Jenkin, & Zikovitz, 2000; Nardini, et al., 2008;

Sun, Campos, & Chan, 2004). By using a cue conflict paradigm, the goal was to determine whether participant responses are closer to the value indicated by the more reliable cue. Although these studies were all motivated by theories of optimal cue combination, they did not sufficiently predict how judgments are based on the reliability of information. As in the current study, Ellard and Shaughnessy (2003) compared the use of visual landmark and walking information for judging the locations of targets. Participants viewed targets at distances of 4 to 10 meters and then walked to them blindfolded. In a few trials, the viewed and walked distances conflicted (one was 6 m and the other 8 m). Participants then reproduced the perceived target distance by walking while blindfolded. For the conflict trials, responses deviated towards the average of the distances indicated by the visual and walking cues. This study calculated the weighting of cues post-hoc based on the relative distance of the response from the visual and walked distances; they did not calculate weights based on the variability of the estimates with only visual or walking information.

Thus far to our knowledge, only the study by Nardini, et al. (2008) made predictions based on the reliability of the cues. Participants learned the locations of several objects in a room by viewing surrounding landmarks and by walking to the objects sequentially. Participants then estimated the location of the first object by walking to its perceived location. In the conflict condition, the landmarks were rotated by fifteen degrees before participants responded. The results indicated that adults integrated landmark and path integration according to their reliabilities, measured when only landmarks or walking information was available, but children alternated between these sources of information. Despite these findings, this study did not test the key prediction that localization estimates change if the reliability of sensory information is altered. In psychophysical studies of low-level cue combination, reliability is manipulated by introducing noise into one or several sensory cues. Increased noise reduces the reliability of the information, causing estimates to be more variable. The current study manipulates visual reliability by having participants wear blur foils that reduce resolution and contrast.

Experiment 1

In the first experiment, we tested the prediction that humans integrate visual landmarks and path integration to determine their position in a hallway based on the reliability of these cues. When the reliability of visual landmark information is reduced by blur, we anticipate that participants will increase their reliance on path integration.

Methods

Participants

Thirteen normally-sighted observers (mean age = 19.7, 8 females/5 males) participated in this study. Participants were compensated monetarily or with extra credit in their psychology course.

Apparatus

Participants were tested in a building hallway approximately 15 meters in length under full lighting (Figure 4.3A). The doors in the hallway remained closed during testing.

The visual targets were red, high-intensity Light Emitting Diodes (LEDs) embedded in wooden sticks lying on the floor and placed every 0.5 meter down the length of the hallway. Each stick contained two LEDs that were turned on individually with varying brightness. The LEDs were controlled via a serial port connection to a laptop.

Participants viewed targets while wearing monocular goggles that were either clear or blurry. The blur goggles were made using Bangerter Occlusion foils placed on the surface of the goggles. The blur foils produced an average logMAR acuity of 1.60 (Snellen acuity of approximately 20/800) and log contrast sensitivity of 0.23. The fellow eye was occluded. The clear goggles were the same as the blur goggles, but without the blur foils. When viewing the hallway through the blur goggles, it was not possible to see the textures on the floor and wall, nor the doors that were farther down the hallway (Figure 4.3B).

Some trials involved walking through the hallway while blindfolded. To prevent veering during these trials, participants held onto a steel cable that was strung along the length of the hallway. The cable was tight enough to prevent any slack from being a distance cue. Participants used tactile markers attached to both ends of the cable to stop before the hallway ended.

During walking trials, a laser range finder was used to measure the distance traveled by participants. The laser range finder was connected to the laptop via a Bluetooth connection, and was used to measure starting and stopping distances. Participants were also stopped at specified distances by a program that used a Kalman filter algorithm to predict the time needed for participants to walk the required distance. The Kalman filter estimated walking time by combining a model of the individual's walking speed with continuous observations of the participant's location acquired by the laser range finder. On average, participants were stopped within +/- 23 centimeters of the desired distance. The laser range finder and Kalman filter were used to reduce the time required to run each trial. Furthermore, we could stop participants at locations without visual markers. Participants wore a white cardboard on their back, which provided a reflective surface for the laser. They also wore noise-reducing headphones to prevent the use of auditory cues to estimate location, but could hear instructions from the experimenter via radio.

In all sessions, participants made location estimates on a tactile map that depicted the doors and intersections in the test hallway (Figure 4.3C). Participants indicated locations while blindfolded by moving a slider across the map that was also attached to a metric ruler. The experimenter recorded responses as the location of the slider on the ruler to the nearest millimeter. We used a tactile map so that the response modality was independent of the stimulus modalities (vision and walking) and did not bias how the cues were weighted in the response. Participants were trained with the tactile map prior to testing (see below).

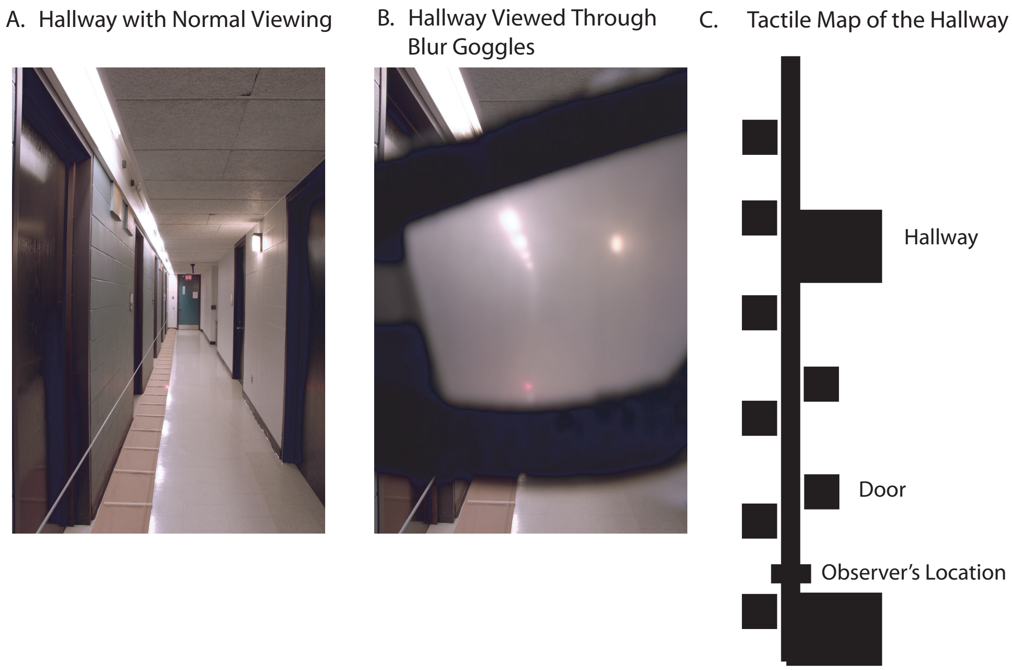

Figure 4.3. Images of the hallway with (A) normal viewing, (B) viewed through the blur goggles, and (C) as represented by the tactile map.

Tasks

All participants were tested in three localization tasks: visual estimation, walking estimation, and combined estimation (visual and walking). The visual and walking estimation tasks provided data on single cue reliability, which were used to make predictions for the combined task.

Visual Estimation

Participants stood at one end of the test hallway and judged the locations of targets (a single LED on a stick) viewed through the clear and blurry goggles. The brightness and the lateral position (the right versus left LED on each stick) of the targets were randomized for each trial to minimize depth information from intensity or visual angle. We set the range of brightness for both viewing conditions based on detection performance in pilot tests and made sure participants could detect the targets with blur at the dimmest setting (see below).

For each trial, participants viewed the target after hearing an auditory cue. They were allowed to look at the target as long as they needed to obtain a good idea of its location. Then, they pulled a blindfold down over the goggles, and indicated on the tactile map the location of the target. The target distances were 5, 7, 9 and 11 meters. Participants performed ten trials at each target distance, plus an additional six trials at alternate distances, to prevent memorization, for a total of forty-six trials in each viewing condition. Participants were standing while performing the task.

Walking Estimation

In this task, participants walked a distance determined by the experimenter while blindfolded and then indicated on the tactile map their location in the hallway. Before walking, participants lifted the blindfold to look at the hallway, but no LED targets were turned on. The purpose of the hallway view was to refresh participants' visual memory of the environment on each trial. Some argue that when humans walk without visual information, location is determined by updating movement with respect to the imagined environment (Rieser & Pick, 2002). Because participants viewed the hallway before walking in each trial of the combined task (see below), we allowed the same information in the walking estimation task.

After viewing the hallway, participants wore the blindfold and began walking. After hearing another auditory signal, participants stopped walking and indicated their perceived location on the tactile map before walking back to the start. Participants performed 40 trials in each viewing condition, ten trials for each target distance (5, 7, 9, and 11 m).

Combined Estimation

In the combined estimation task, participants judged their location after viewing a target and then walking a specified distance. The visual targets were located at 7 and 9 meters. Participants walked to locations that were the same as the visual target (conflict of 0 m) or differed by +/- 0.25, 0.5, or 0.75 meters. Pilot tests revealed that participants were highly likely to detect conflicts of 1 meter or greater. As in the walking estimation task, participants indicated their location on the tactile map after an auditory cue signaled for them to stop walking. As they walked back to the starting location, they also reported whether they thought the location of the viewed target and the walked location were the same or different. Participants also rated their confidence of the same/different locations judgment, but these ratings will not be analyzed in this paper. Participants performed five trials for each visual target and conflict for a total of 70 trials for each viewing condition.

Procedure

In each task, participants were tested in two viewing conditions: no blur and blur. Half of the participants performed the no blur condition first, and the other half performed the blur condition first. The experiment was conducted in four sessions, each lasting about one and a half hours, in the following order: 1) visual estimation task in both viewing conditions, 2) walking estimation task in both viewing conditions, 3) combined task in one viewing condition, 4) combined task in the remaining viewing condition.

Training

Participants first performed a set of visual tests to evaluate their visual ability with the clear and blurry goggles. They performed the tests and the rest of the experiment with their dominant eye determined with the Miles test (Miles, 1930). Monocular visual acuity was measured with both the clear and blurry goggles using a Lighthouse Distance Acuity test. Contrast sensitivity was measured with the Pelli-Robson Contrast Sensitivity chart.

Participants were also tested on their ability to see the LED targets with the blurry goggles. Forced-choice testing confirmed that all subjects could detect the LED targets under all conditions with at least 90% accuracy.

In the first session, participants were also trained to use the tactile map. The experimenter explained the features and scaling of the map, and how to use the pointer. Participants studied the map and the hallway, ensuring they understood the features in the environment prior to the blur conditions. The experimenter then walked down the hallway with the subject, stopped at random locations, and asked the participant to indicate their location while looking at the map. Participants then practiced making responses on the map while blindfolded.

At the beginning of the walking and combined estimation tasks, participants practiced walking blindfolded through the hallway while holding on to the cable. They were instructed to walk back and forth in the hallway, stopping themselves at the tactile markers, until they felt comfortable and were able to walk at a normal pace. The experimenter also assured participants that they would be warned of any people or obstructions that appeared in the hallway. The experimenter also measured participants' walking velocity and step length for each session. The velocity information was entered into the Kalman filter algorithm to determine when to stop the participant on each trial.

Data Analysis

This experiment tested two main predictions: 1) observers integrate visual and walking information to determine their position in the hallway only when they perceive the information as congruent rather than conflicting, and 2) when these cues are integrated, observers weigh walking information more as the reliability of visual information decreases.

First, we used the visual and walking estimation tasks to measure the reliability of these cues for localization. Reliability is equivalent to the inverse of the variability of the estimates ($1/\sigma^2$). Visual and walking variability were measured by first fitting separate lines to the responses in each task and for each condition (normal and blurry viewing). We accounted for individual biases to prevent inflation of variability from between subject differences (see Appendix). Variability (σ_v and σ_w) was calculated as the root mean square of the residuals of the best fit lines. Using these variability measures, we computed the predicted weights (w_v and w_w) of visual and walking information for the combined task (Equation 4.3).

To evaluate whether integration depended on the congruence of information, we separated responses in the combination task according to whether participants perceived a conflict between the viewed and walked locations. For both sets of data, we tested whether there was a qualitative shift in the reliance on walking information between the normal and blurry viewing conditions. We anticipated that an increase in reliance on walking information would be demonstrated by an increase in the slope of the responses when plotted against the distance walked on each trial (Figure 4.4). For example, if participants were only relying on visual information to judge their location (weight of walking information is 0), then their responses should always equal the distances of the visual targets at 7 or 9 meters; responses should not change with the walked distance (slope of 0 if the 7 and 9 meter trials are plotted separately). However, if participants only relied on walking information (weight of walking information is 1), then their responses should equal the walked distance (slope of 1). Accordingly, the slope of the plot of estimated versus actual walked distance is equivalent to the weight of the walking information.

For trials in which participants did not perceive a conflict between the viewed and walked locations (congruent trials), we predicted an increase in the slope of the responses in the blurry viewing condition compared to the normal viewing condition, corresponding to an increased reliance on walking information. For trials in which a conflict was perceived (incongruent trials), we predicted that the slopes of the responses would be near 1 because participants would only rely on walking information to judge their location in the hallway. For both congruent and incongruent trials, we tested for changes in slope at each visual target distance (7 and 9 meters) by fitting separate lines to the data from the normal and blurry viewing conditions. We then performed contrasts between the slopes of the fitted lines (7 m with normal viewing versus 7 m with blurry viewing, and 9 m with normal viewing versus 9 m with blurry viewing). Significant contrasts suggested that walking information was weighted differently in the two viewing conditions. This analysis was conducted on the data grouped across participants as well as individual data.

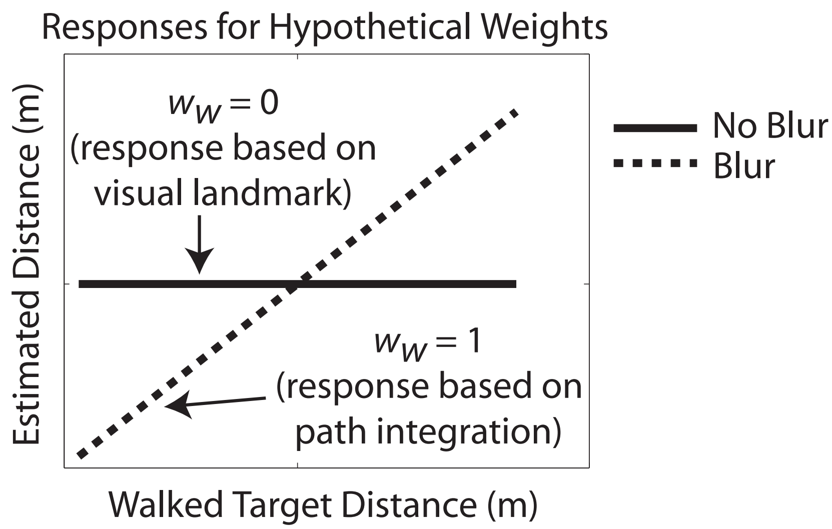

Figure 4.4. Expected responses in the combined task if walking information has a weight of 0 in the normal viewing condition and a weight of 1 in the blurry viewing condition.

We evaluated the quantitative predictions of the model by comparing the predicted weights from the optimal combination rule to the weights obtained from participants' responses. Since the weight of walking information is equivalent to the slope of the line fitted to the responses (described above), we computed 95% confidence intervals on these slope coefficients. Our quantitative prediction was supported if the confidence intervals included the predicted weights. This analysis was only conducted on the data grouped across participants; individual data did not provide accurate estimates of reliability because there were an insufficient number of data points per participant.

We used robust linear models (with a bisquare estimator) to obtain the best fits of the data in all three tasks. The purpose of using robust fitting methods was to reduce the influence of outlying data points since they increased the variability of the estimates. Robust fits are derived by an iteratively reweighted least squares algorithm that gives lower weights to points that are not fit well. We used these weights when calculating the variability of estimates, which was the root mean of the weighted squared residuals.

<u>Results</u>

Reliability of Visual and Walking Information

Figure 4.5 depicts responses for all participants in the visual and walking estimation tasks. The variability of the estimates for each condition is displayed in Table 4.1 for the distances tested in the combined task. As expected, visual estimates were more variable with blurry viewing compared to normal viewing. Viewing condition did not have as great an impact on walking estimation. The variability of the single cue estimates (σ_v and σ_w) were used to predict the weights (w_w) of these cues in the combined task. The predicted weights for walking information, also shown in Table 4.1, indicate that walking information should have a greater weighting with blurry vision compared to normal vision when visual and walking information are integrated.

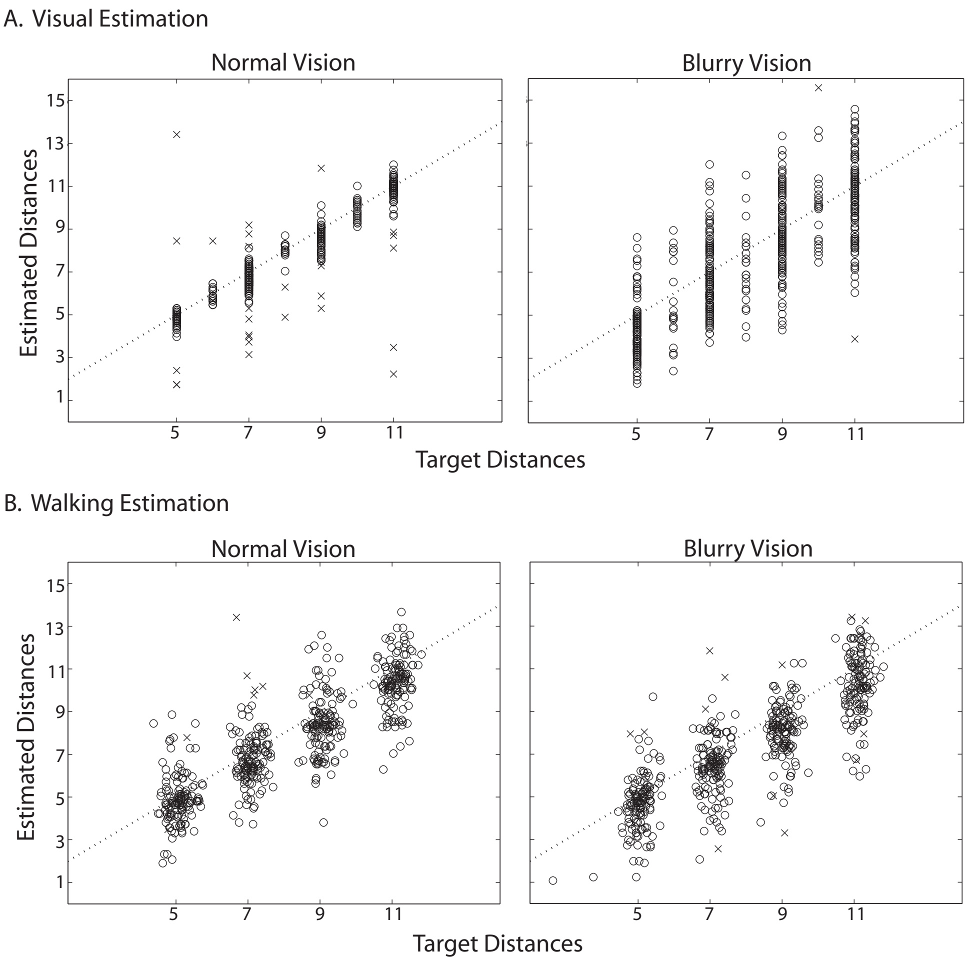

Figure 4.5. (A) Visual and (B) walking estimates in the normal and blurry viewing conditions compiled across participants. Estimates determined to be outliers by robust fitting are marked by an ‘x.’

	Variability of Visual Estimation (σ_v)	Variability of Walking Estimation (σ_w)	Predicted Weights of Walking Information $w_w = \frac{1/\sigma_w^2}{1/\sigma_v^2 + 1/\sigma_w^2}$
Normal Viewing	7 m: 0.031 m 9 m: 0.028 m	7 m: 0.068 m 9 m: 0.084 m	7 m: 0.17 9 m: 0.10
Blurry Viewing	7 m: 0.116 m 9 m: 0.133 m	7 m: 0.081 m 9 m: 0.074 m	7 m: 0.67 9 m: 0.76

Table 4.1. Variability (root mean square errors) and predicted weights of walking information computed from the visual and walking estimation tasks for both viewing conditions at 7 and 9 meters.

Analysis of Grouped Data

We separately analyzed trials in which participants judged the visual and walked locations to be the same versus different. On average, participants judged the viewed and walked locations to be the same on approximately half the trials in both viewing conditions (55% of trials in the normal viewing condition and 51% of trials in the blurry viewing condition).

Figure 4.6 shows responses for all participants for both trial types (locations judged to be same or different), and the best fit lines for four conditions (visual target (7 m and 9 m) x viewing condition (normal and blurry)). For trials in which viewed and walked locations were perceived as the same, contrasts revealed that walking information was weighted significantly higher in the blur condition compared to the normal vision condition for visual targets at 9 m ($p < 0.001$). There was no significant difference between the weights for the 7 m visual target ($p = 0.436$). For trials in which a discrepancy was perceived between the visual and walked locations, there were no significant changes in weighting between the normal and blurry viewing conditions.

To compare the observed and predicted weights, we computed 95% confidence intervals for the slopes of the best fit lines in each condition (Figure 4.7). For trials in which the viewed and walked locations were judged to be the same, only the observed weight of the 7 m blurry viewing condition matched the predicted weight. For the other conditions, the observed weighting of walking information was higher than predicted. When viewed and walked locations were perceived as different, the weighting of walking information was approximately 1 in all conditions. This confirmed that participants did not integrate the two sources of information when they were perceived as discrepant, but instead only used walking information.

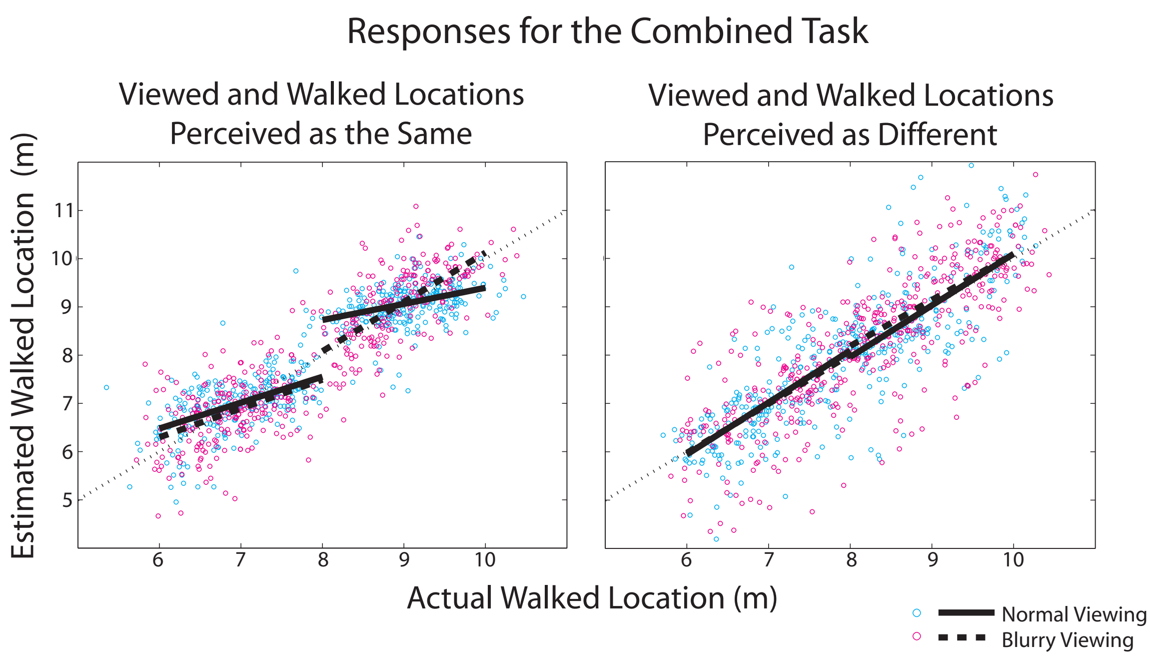

Figure 4.6. Participant responses in the combination task. These data were normalized by subtracting off individual constant biases from each participants' responses.

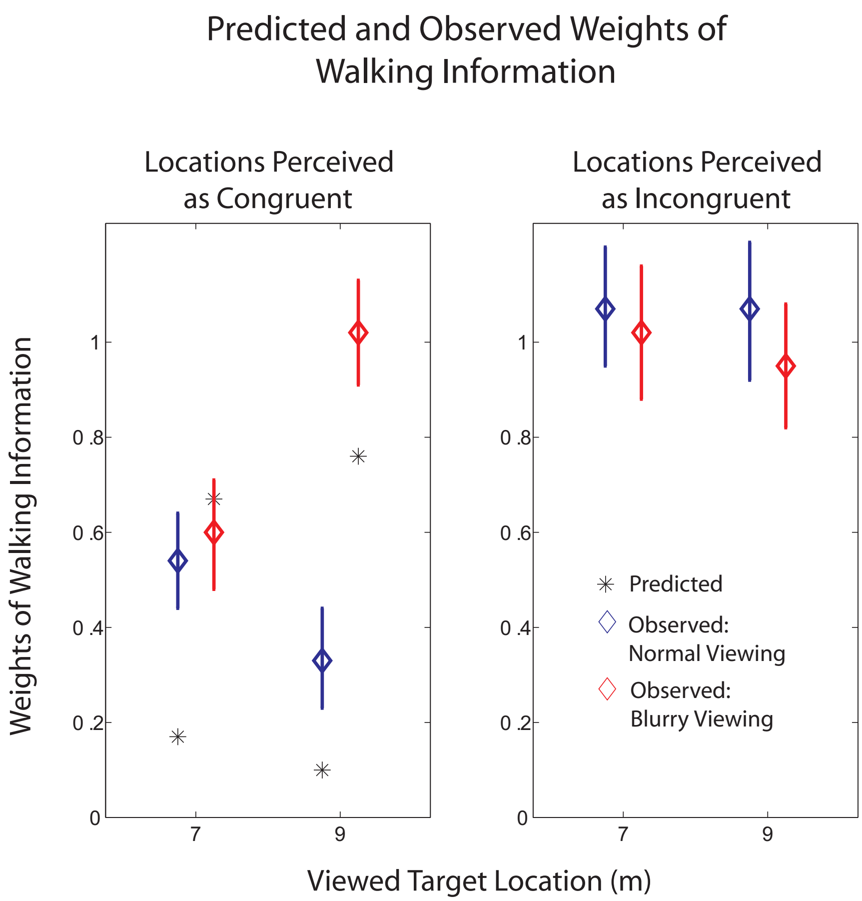

Figure 4.7. Predicted (asterisks) and observed (diamonds) weights of walking information for congruent and incongruent trials. Error bars represent 95% confidence intervals for the observed weights.

Analysis of Individual Data

We also conducted statistical tests of the qualitative predictions of individual estimates in the cue combination task. Data from two representative participants when they did not perceive conflicts are shown in Figure 4.8. For the 7 meter distance when participants did not perceive a conflict, four subjects showed significant increases in the weighting of walking information and four others showed trends in the predicted direction. In the 9 meter condition, ten subjects had significant increases in the weighting of walking information and the other three displayed trends in the predicted direction. When participants did perceive a conflict, only 3 subjects showed significant increases in the weighting of walking information for both distances. Thus, the analysis of individual estimates also shows that participants performed according to the predictions of the cue combination model at the 9 meter distance, but less so for the 7 meter distance, when they did not perceive a conflict, but relied only on walking information when they did perceive a conflict.

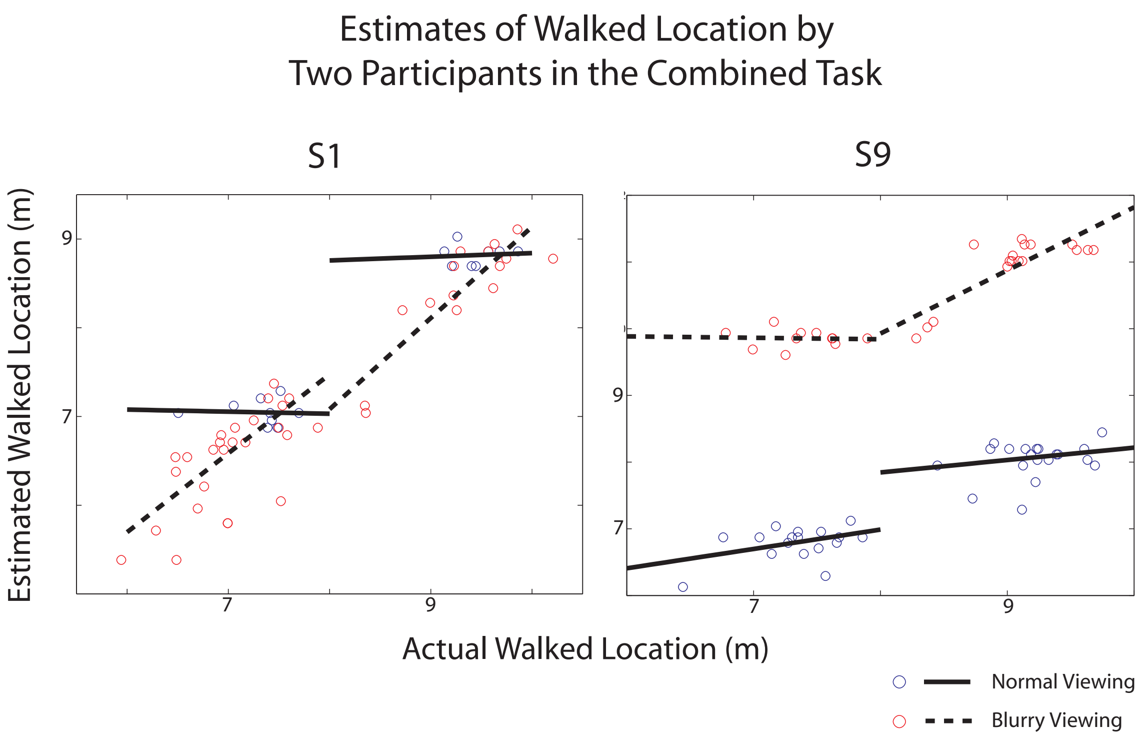

Figure 4.8. Responses of participants S1 and S9 for trials for which the visual and walked locations were judged to be the same.

<u>Discussion</u>

Our results show that participants combined visual estimates of landmark positions and path integration depending on the reliability and congruency of the information. When the viewed and walked locations were perceived to be the same, participants in some cases weighed these cues according to their reliability. For distances around 9 meters, participants increased their weighting of walking information as the reliability of visual information decreased with blur. Participants weighed walking information more than predicted in all conditions except for visual targets at 7 m viewed with blur; therefore, we did not observe a change in the reliance on walking information between the 7 m normal versus blurry viewing conditions. Furthermore, participants only relied on walking information to determine their location when they perceived the viewed location to be different from the walked location.

One possible explanation for the greater reliance on walking information is that visual reliability decreases over time due to memory decay. Accordingly, our estimates of visual reliability as measured in the single cue visual task would inaccurately represent this information as being more reliable than it actually is. We tested this possibility in Experiment 2 in which participants made visual estimates after a time delay.

Experiment 2

The cue combination task in Experiment 1 required participants to learn a cognitive map of the environment and to maintain the locations of visual landmarks in memory while walking. Participants then compared the walked location to the location of the remembered landmark. How well is the location of the visual landmark preserved during the walk? Previous research on visual memory decay suggests that either visual memory accumulates noise over time (Kinchla & Smyzer, 1967) or decays in a more deterministic fashion (Gold, et al., 2005). The goal of this experiment was to investigate if the reliability of visual information for localization decays over time.

Method

Participants

We tested six normally-sighted participants (mean age = 22, 2 females/4 males) who were compensated monetarily or with extra credit in their psychology course.

Apparatus

The test environment and the other materials were the same as in Experiment 1. Because participants did not walk blindfolded in this experiment, we did not use the laser range finder and participants did not wear the white board or headphones.

Procedure

Participants performed the visual estimation task, described in Experiment 1, in four conditions: two viewing conditions (normal versus blurry viewing) with two delay conditions (no delay versus delay). In the delay condition, participants waited for a set amount of time between viewing the target and responding on the tactile map. The delay corresponded to the amount of time the participant needed to walk to the target. The order of the conditions was randomized for each participant.

First, participants went through the same pre-experiment training as in Experiment 1. We measured each participant's walking velocity (mean = 1.03 m/s) to calculate response delays. The delay was determined by estimating the time required for the participant to walk to the target blindfolded (time delay (s) = target distance (m) / walking velocity (m/s)). Therefore, time delays on average ranged from 4.85 seconds for targets at 5 meters to 10.68 seconds for targets at 11 meters. Participants were then tested in the visual estimation task in the four conditions. The no delay condition was the same as the visual estimation task in Experiment 1. In the delay condition, participants also viewed the target after hearing an auditory cue. After obtaining a good view of the target, participants wore the blindfold and waited until they heard another auditory cue. Then the participant indicated the location of the target on the tactile map. As in Experiment 1, participants performed 46 trials in each condition, 10 trials at 5, 7, 9, and 11 meters and 6 trials at alternate distances to prevent memorization.

Data Analysis

Robust linear models were fit to each participant's data in each of the four conditions (no delay versus delay, normal and blurry viewing). We computed the variability of estimates as the root mean of the weighted squared residuals across all target distances. To test the effect of delay, we performed paired t-tests for each viewing condition comparing the variability of estimates between the no delay and delay conditions.

Results

The variability of estimates in the four conditions, averaged across participants, were as follows: normal viewing - no delay: 0.023 m, normal viewing - delay: 0.022 m, blurry viewing – no delay: 0.093 m, blurry viewing – delay: 0.109 m. Paired t-tests revealed no significant differences between the no delay and delay conditions in both viewing conditions.

Discussion

This experiment revealed that the reliability of visual information for target locations did not decrease over the time needed to walk the same distance. We can conclude that the delay between viewing the target and responding in the combination task of Experiment 1 does not account for the greater than predicted reliance on walking information. Although longer delays or distractor tasks could be more detrimental to the visual memory of targets, previous studies show that memory for landmarks is highly robust over time, and remains accurate even a year after learning (Stankiewicz & Kalia, 2007). The results of Experiment 2 corroborate these previous findings.

General Discussion

Humans combine information from visual landmarks and path integration depending on the congruency and reliability of these cues. Participants combined both sources of information only when they judged the viewed and walked locations to be the same. Otherwise, they only used walking information to determine their location in the hallway. Furthermore, participants accounted for the reliability of information when combining landmarks with path integration. For one of the two target distances (9 m), participants used information obtained by walking more as the reliability of visual information decreased. For the shorter distances (around 7 meters), participants integrated the cues, but did not shift their reliance based on the quality of the visual information. Instead, they seemed to average the estimates provided by vision and walking, regardless of the viewing condition. These results illuminate the conditions required for humans to integrate multiple sources of information when determining their location.

In Experiment 1, we found that participants sometimes relied more on walking information than predicted. Experiment 2 explored whether the greater reliance on walking information was due to the delay between viewing the target and responding. We specifically tested if the memory of visual landmarks became noisy during the time participants walked to the targets. The results of this experiment revealed that the reliability of visual information remained consistent during short time intervals. Previous studies found that visual landmark information remains robust over time, more so than path integration (Zeigler & Wehner, 1997; Stankiewicz & Kalia, 2007). Perhaps the integration of visual landmarks into a cognitive map allows this information to be stable over long time delays.

Another possible explanation for the greater than expected reliance on path integration is the order in which viewed and walked locations were presented. Unlike typical cue conflict studies, participants in our experiment were presented with information sequentially- they first viewed a target and then walked to it. Ellard and Shaughnessy (2003) found that the order of presentation influenced how visual and walking information were combined to estimate distances. When the walking information was presented last, participants weighed both cues about equally. When the visual information was presented last, the visual weighting was significantly greater than 0.5. Therefore, it is possible that the recency of information influenced how participants weighed walking information in our experiment. This hypothesis will be tested in future experiments.

The exploration of cue congruency has important theoretical implications for the study of how the perceptual system integrates information from multiple sources. A recent branch of research explores the problem of causal inference and its relationship to sensory integration. The hypothesis is that the perceptual system should treat noticeably discrepant cues as coming from different sources, and this information should not be integrated (Körding, et al., 2007). Although the underlying computation is the same (separation versus integration of information), determining the congruence of information is a different problem. In this study, visual and walking information clearly come from two different sources that are separated temporally. The question of interest is whether both sources provide information that is relevant to the task. We measure relevance as the spatial congruence of the position estimates provided by vision and walking. The ability to identify and integrate relevant sources of information for a task may develop with age (Nardini, et al., 2008).

Also unlike previous studies on cue integration for navigation, we altered the reliability of information to determine if cue weights change accordingly. In addition to confirming the predictions of the statistical model, manipulating the visual factors of acuity and contrast has clear implications for understanding how individuals with low vision combine sensory information to estimate locations. People with visual impairment do not express increased sensitivity to non-visual information obtained by walking, as demonstrated by studies of path integration with blind individuals (Loomis, et al., 1993). Yet, it may be that people with low vision adjust how they weight visual and non-visual cues to obtain the most optimal perceptual estimates given the available information. Specifically, reliance on non-visual information should increase with the severity of the visual impairment because visual reliability decreases. Exploring whether people with visual impairment optimally integrate residual vision with other sensory information can be a useful test of the effectiveness of mobility training.

In conclusion, this study provides evidence that humans can optimally integrate information from visual landmarks and path integration to determine their location in an environment. Furthermore, integration is dependent on the congruence of the information- humans will only incorporate landmarks that fall within the range of locations specified by path integration. Accordingly, information integration for navigation is influenced by the properties of underlying sensory information.

Appendix: Accounting for Individual Differences in Localization Estimates

Measuring Reliability

To determine the implications for grouping data across participants, we first compared single cue estimates between participants to evaluate the extent of individual variations. Robust linear models were fit to the estimates for each participant for each task (visual and walking estimation) and condition (normal and blurry viewing). The coefficients (β_0 and β_1) obtained from these fits were compared across individuals using contrasts. Perfect estimates of viewed and walked locations would yield an intercept (β_0) of 0 and a slope (β_1) of 1. Slopes that deviate significantly from 1 indicate a scaling bias. Intercepts that deviate from 0 can be due to constant (overall tendency to under or overestimate distances) and scaling biases.

The results of the contrasts between individual fits, summarized in Table 4A.1, indicate that visual estimates were comparable across participants ($p > 0.05$) whereas walking estimates varied significantly across participants, particularly the fitted slopes of these estimates ($p < 0.01$). Accordingly, we estimated visual variability by fitting the pooled visual estimates, and computed variance as the root mean of the weighted squared residuals for a particular distance. To compute walking variability, we fit the walking estimates separately for each individual and then computed root mean square error across participants to obtain a single estimate of variability in each viewing condition. We calculated variability separately for target distances of 7 and 9 meters, since these were the distances used in the combined task.

Task - Condition	Coefficient	Range	P value of Contrast
Visual – No Blur	β_0	-0.49 – 0.07	0.064
	β_1	0.98 – 1.03	0.343
Visual – Blur	β_0	-3.41 - 1.40	0.431
	β_1	0.83 - 1.26	0.816
Walking – No Blur	β_0	-1.09 - 1.44	0.150
	β_1	0.81 - 1.10	0.008
Walking - Blur	β_0	-1.79 - 1.20	0.028
	β_1	0.73 - 1.21	0.006

Table 4A.1. Coefficients of individual fits of single cue estimates and the p-value of contrasts comparing the coefficients across individuals.

Individual Differences in the Combined Task

For the analysis of the grouped data, we pooled the estimates across participants after normalizing for individual constant biases (the average tendency to overestimate or underestimate). We removed constant biases since we were interested in changes in the fitted slopes of the estimates (Figure 4.4). From the analysis in Table 4A.1, we acknowledge that individual variations in scaling are also likely in the combined task, which would affect the slopes of estimates in addition to changes in cue weightings. To address this issue, we also analyzed the individual responses in the combined task.

Bibliography

Chance, S.S., Gaunet, F., Beall, A.C., & Loomis, J.M. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence: Teleoperators and Virtual Environments, 7(2),* 168-178.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition, 23,* 149-178.

Cheng, K., & Newcombe, N.S. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic Bulletin & Review, 12(1),* 1-23.

Collett, T.S., Graham, P., & Durier, V. (2003). Route learning by insects. *Current Opinion in Neurobiology, 13,* 718-725.

Dijkstra, E.W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik, 1*, 269-271.

Ellard, C. G., & Shaughnessy, S. C. (2003). A comparison of visual and nonvisual sensory inputs to walked distance in a blind-walking task. *Perception, 32(5),* 567-578.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Science, 8(4),* 162-169.

Etienne, A.S., Maurer, R., Boulens, V., Levy, A., & Rowe, T. (2004). Resetting the path integrator: A basic condition for route-based navigation. *The Journal of Experimental Biology, 207,* 1491-1508.

Etienne, A., Maurer, R., & Seguinot, V. (1996). Path integration in mammals and its interaction with visual landmarks. *The Journal of Experimental Biology, 199,* 201-209.

Etienne, A. S., Teroni, V., Hurni, C., & Portenier, V. (1990). The effect of a single light cue on homing behaviour of the golden hamster. *Animal Behavior, 39*, 17–41.

Eye Disease Prevalence Research Group. (2004). Causes and prevalence of visual impairment among adults in the United States. *Archives of Ophthalmology, 122*, 477-485.

Farrell, M.J., Arnold, P., Pettifer, S., Adams, J., Graham, T., & MacManamon, M. (2003). Transfer of route learning from virtual to real environments. *Journal of Experimental Psychology: Applied, 9(4),* 219-227.

Gallistel, C.R. (1990). *The Organization of Learning.* Cambridge, MA: MIT Press.

Gibson, J.J. (1950). *The Perception of the Visual World*. Boston: Houghton Mifflin Company.

Giudice, N. A. (2004). Navigating novel environments: A comparison of verbal and visual learning. (Doctoral dissertation, University of Minnesota, 2004). *Dissertation Abstracts International B, 65,* 6064.

Giudice, N. A., Bakdash, J. Z., & Legge, G. E. (2007). Wayfinding with words: Spatial learning and navigation using dynamically-updated verbal descriptions. *Psychological Research, 71(3),* 347-358.

Giudice, N.A., & Legge, G.E. (2008). Blind navigation and the role of technology. In A. Helal, M. Mokhtari, & B. Abdulrazak (eds.), *The Engineering Handbook on Smart Technology for Aging, Disability and Independence*. Hoboken, NJ: John Wiley & Sons, Inc.

Gold, J.M., Murray, R.F., Sekuler, A.B., Bennett, P.J., & Sekuler, R. (2005). Visual memory decay is deterministic. *Psychological Science, 16(10),* 769-774.

Golledge, R.G. (1999). Human wayfinding and cognitive maps. In R.G. Golledge (ed.), *Wayfinding Behavior: Cognitive Mapping and other Spatial Processes.* Baltimore: The Johns Hopkins University Press.

Golledge, R. G., Marston, J. R., Loomis, J. M., & Klatzky, R. L. (2004). Stated preferences for components of a Personal Guidance System for nonvisual navigation. *Journal of Visual Impairment & Blindness, 98,* 135-147.

Gouteux, S., & Thinus-Blanc, C. (2001). Rhesus monkeys use geometric and non-geometric information during a reorientation task. *Journal of Experimental Psychology: General, 130(3),* 505-519.

Green, C.S. & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423,* 534-537.

Harris, L. R., Jenkin, M., & Zikovitz, D. C. (2000). Visual and non-visual cues in the perception of linear self-motion. *Experimental Brain Research, 135(1),* 12-21.

Hermer, L., & Spelke, E. (1996). Modularity and development: the case of spatial reorientation. *Cognition, 61,* 195-232.

Holmes, E., Jansson, G., & Jansson, A. (1996). Exploring auditorily enhanced tactile maps for travel in new environments. *New Technologies in the Education of the Visually Handicapped, 237,* 191-196.

Kalia, A.A., Schrater, P.R., Legge, G.E., & Kallie, C.S. (2008, May). Estimating absolute distances with blurred vision. Poster presented at the Annual Meeting of the Vision Sciences Society, Naples, Florida.

Kelly, D., Spetch, M., & Heth, C. (1998). Pigeons' (Columba livia) encoding of geometric and featural properties of a spatial environment. *Journal of Comparative Psychology, 112(3),* 259-269.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as bayesian inference. *Annual Review of Psychology, 55,* 271-304.

Kinchla, R.A., & Smyzer, F. (1967). A diffusion model of perceptual memory. *Perception and Psychophysics, 2,* 219-229.

Kirasic, K. C., Allen, G.L, Dobson, S.H., & Binder, K.S. (1996). Aging, cognitive resources, and declarative learning. *Psychology and Aging, 11(4),* 658-670.

Klatzky, R.L., Loomis, J.M., Beall, A.C., Chance, S.E., & Golledge, R.G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science, 9(4),* 293-298.

Körding, K.P., Beierholm, U., Ma, W.J., Quartz, S., Tenenbaum, J.B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE, 2(9),* e943. doi: 10.1371/journal.pone.0000943.

Kuipers, B., Tecuci, D. G., & Stankiewicz, B. J. (2003). The skeleton in the cognitive map: A computational and empirical exploration. *Journal of Environment and Behavior, 35,* 81–106.

Kuyk, T., & Elliott, J.L. (1999). Visual factors and mobility in persons with age-related macular degeneration. *Journal of Rehabilitation Research & Development, 36(4),* 303-312.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35(3)*, 389-412.

Lessels, S., & Ruddle, R. A. (2004). Changes in navigational behaviour produced by a wide field of view and a high fidelity visual scene. *Proceedings of the 10th Eurographics Symposium on Virtual Environments (EGVE'04)*, 71-78.

Lessels, S., & Ruddle, R. A. (2005). Movement around real and virtual cluttered environments. *Presence: Teleoperators and Virtual Environments, 14(5),* 580-596.

Loomis, J.M., Golledge, R.G., & Klatzky, R.L. (1998). Navigation system for the blind: Auditory display modes and guidance. *Presence: Teleoperators and Virtual Environments, 7(2),* 193-200.

Loomis, J.M., Golledge, R.G., & Klatzky, R.L. (2001). GPS-based navigation systems for the visually impaired. In W. Barfield, & T. Caudell (eds.), *Fundamentals of Wearable Computers and Augmented Reality*. Mahwah, NJ: Lawrence Erlbaum Associates.

Loomis, J.M., Klatzky, R.L., Avraamides, M., Lippa, Y., & Golledge, R.G. (2007). Functional equivalence of spatial images produced by perception and spatial language. In F. Mast, & L. Jäncke (eds.), *Spatial Processing in Navigation, Imagery, and Perception*. New York: Springer.

Loomis, J., Klatzky, R., Golledge, R., & Philbeck, J. (1999). Human navigation by path integration. In Golledge, R. (ed.), *Wayfinding: Cognitive Mapping and other Spatial Processes*. Baltimore: Johns Hopkins.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General, 122(1),* 73-91.

Loomis, J. M., Lippa, Y., Klatzky, R. L., & Golledge, R. G. (2002). Spatial updating of locations specified by 3-D sound and spatial language. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 335-345.

Ludt, R., & Goodrich, G.L. (2002). Change in visual perceptual detection distances for low vision travelers as a result of dynamic visual assessment and training. *Journal of Visual Impairment & Blindness, 96(1),* 7-21.

Lund, R., & Watson, G.R. (1997). *The CCTV Book: Habilitation and Rehabilitation with Closed Circuit Television Systems.* Frolund, Norway: Synsforum.

Mamassian, P., Landy, M., & Maloney, L. T. (2002). Bayesian modeling of visual perception. In R. P. N. Rao, B. A. Olshausen, & M. S. Lewicki (eds.), *Probabilistic Models of the Brain: Perception and Neural Function*. Cambridge, MA: MIT Press.

Mallot, H. A., & Gillner, S. (2000). Route navigating without place recognition: What is recognised in recognition-triggered responses? *Perception, 29,* 43–55.

Marron, J.A. & Bailey, I.L. (1982). Visual factors and orientation-mobility performance. *American Journal of Optometry and Physiological Optics, 59(5),* 413-426.

Mason, S. J. (2002). Indoor way-finding with restricted visual fields. Unpublished Honors thesis, University of Minnesota, Minneapolis.

Mason, S.J., Legge, G.E., & Kallie, C.S. (2005). Variability in the length and frequency of steps of sighted and visually impaired walkers. *Journal of Visual Impairment & Blindness, 99,* 741-54.

Moffat, S. D., & Resnick, S.M. (2002). Effects of age on virtual environment place navigation and allocentric cognitive mapping. *Behavioral Neuroscience, 116(5),* 851-859.

Moffat, S. D., Zonderman, A.B., & Resnick, S.M. (2001). Age differences in spatial memory in a virtual environment navigation task. *Neurobiology of Aging, 22(5),* 787-796.

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology, 18,* 689-693.

Ooi, T.L, Wu, B., & He, Z.J. (2001). Distance determined by the angular declination below the horizon. *Nature, 414,* 197-200.

Palmer, S. E. (1999). *Vision Science: Photons of Phenomenology*. Cambridge, MA: The MIT Press.

Peli, E. (2005). Recognition performance and perceived quality of video enhanced for the visually impaired. *Ophthalmic Physiological Optics, 25(6),* 543-555.

Peli, E., Lee, E., Trempe, C.L., & Buzney, S. (1994). Image enhancement for the visually impaired: The effects of enhancement on face recognition. *Journal of the Optical Society of America, 11(7),* 1929-1939.

Redish, D.A. (1999). *Beyond the Cognitive Map: From Place Cells to Episodic Memory*. Cambridge, MA: MIT Press.

Rieser, J.J., Hill, E.W., Talor, C.R, Bradfield, A., & Rosen, S. (1992). Visual experience, visual field size, and the development of nonvisual sensitivity to the spatial structure of outdoor neighborhoods explored by walking. *Journal of Experimental Psychology: General, 121(2),* 210-221.

Rieser, J. J., & Pick, H. L. (2002). The perception and representation of human locomotion. In W. Prinz & B. Hommel (eds.), *Attention and Performance XIX: Common Mechanisms in Perception and Action.* New York: Oxford University Press.

Rosenzweig, E.S., Redish, A.D., McNaughton, B.L., & Barnes, C.A. (2003). Hippocampal map realignment and spatial learning. *Nature Neuroscience, 6(6),* 609-615.

Ruddle, R. A., & Lessels, S. (2006). For efficient navigational search humans require full physical movement but not a rich visual scene. *Psychological Science, 17,* 460-465.

Ruddle, R.A., Payne, S.A., & Jones, D.M. (1997). Navigating buildings in “desk-top” virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied, 3(2),* 143-159.

Ruddle, R. A., & Peruch, P. (2004). Effects of proprioceptive feedback and environmental characteristics on spatial learning in virtual environments. *International Journal of Human Computer Studies, 60,* 299-326.

Schölkopf, B., & Mallot, H. A. (1995). View-based cognitive mapping and path planning. *Adaptive Behavior, 3,* 311–348.

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*. New York: Academic Press.

Sjölinder, M., K. Hook, K., Nilsson, L., & Andersson, G. (2005). Age differences and the acquisition of spatial knowledge in a three-dimensional environment: Evaluating the use of an overview map as a navigation aid. *International Journal of Human-Computer Studies, 63(6),* 537-564.

Sovrano, V., Bisazza, A., & Vallortigara, G. (2002). Modularity and spatial reorientation in a simple mind: Encoding of geometric and non-geometric properties of a spatial environment by fish. *Cognition, 85,* B51-B59.

Stankiewicz, B.J., & Kalia, A.A. (2007). Acquisition of structural versus object landmark knowledge. *Journal of Experimental Psychology: Human Perception & Performance, 33(2),* 378-390.

Stankiewicz, B.J., Legge, G. E., Mansfield, J. S., & Schlicht, E. J. (2006). Lost in virtual space: Studies in human and ideal spatial navigation. *Journal of Experimental Psychology: Human Perception & Performance, 32(3),* 688-704.

Sun, H. J., Campos, J. L., & Chan, G. S. (2004). Multisensory integration in the estimation of relative path length. *Experimental Brain Research, 154(2)*, 246-254.

Thinus-Blanc, C., & Gaunet, F. (1997). Representation of space in blind persons: Vision as a spatial sense? *Psychological Bulletin, 121(1)*, 20-42.

Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H**.**, Loomis, J. M., & Beall, A. C. (2004). Does the quality of the computer graphics matter when judging distance in visually immersive environments? *Presence: Teleoperators and Virtual Environments, 13,* 560–571.

Thrun, S. (2002). Robotic mapping: A survey. In G. Lakemeyer and B. Nebel (eds.), *Exploring Artificial Intelligence in the New Millenium.* New York: Morgan Kaufmann Publishers.

Tolman, E.C. (1948). Cognitive maps in rats and men. *The Psychological Review, 55(4),* 189-208.

Turano, K. A., Broman, A.T., Bandeen-Roche, K., Munoz, B., Rubin, G.S., & West, S.K. (2004). Association of visual field loss and mobility performance in older adults: Salisbury Eye Evaluation Study. *Optometry & Vision Science, 81(5),* 298-307.

Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology, 13,* 407-433.

Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. *Lecture Notes in Computer Science*, *716,* 14-24.

Ungar, S., Blades, M., & Spencer, C. (1996). The construction of cognitive maps by children with visual impairments. In J. Portugali (ed.), *The Construction of Cognitive Maps*. Dordrecht: Kluwer Academic Publishers.

Vallortigara, G., Zanforlin, M., & Pasti, G. (1990). Geometric modules in animals' spatial representations: A test with chicks (Gallus gallus domesticus). *Journal of Comparative Psychology, 104(3),* 248-254.

van Beers, R. J., Sittig, A. C., & Gon, J. J. D. v. d. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of Neurophysiology, 81(3),* 1355-1364.

Waller, D. (2000). Individual differences in spatial learning from computer-simulated environments. *Journal of Experimental Psychology: Applied, 6(4),* 307-321.

Waller, D., Knapp, D., & Hunt, E. (2001). Spatial representations of virtual mazes: The role of visual fidelity and individual differences. *Human Factors, 43(1),* 147-158.

Waller, D., Loomis, J. M., & Haun, D. B. M. (2004). Body-based senses enhance knowledge of directions in large-scale environments. *Psychonomic Bulletin & Review, 11,* 157-163.

West, S. K., Rubin, G.S., Broman, A.T., Munoz, B., Bandeen-Roche, K., & Turano, K. (2002). How does visual impairment affect performance on tasks of everyday life? *Archives of Ophthalmology, 120,* 774-780.

Wilkniss, S. M., Jones, M.G., Korol, D.L., Gold, P.E., & Manning, C.A. (1997). Age-related differences in an ecologically based study of route learning. *Psychology and Aging, 12(2),* 372-375.

Wilson, I. A., Ikonen, S., Gureviciene, I., McMahan, R.W., Gallagher, M., Eichenbaum, H., et al. (2004). Cognitive aging and the hippocampus: how old rats represent new environments. *Journal of Neuroscience, 24(15),* 3870-3878.

Zeigler, P.E., & Wehner, R. (1997). Time-courses of memory decay in vector-based and landmark-based systems of navigation in desert ants, Cataglyphis fortis. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, *181(1)*, 13-20.