Spring 5-6-2016

Evaluation of Multi-Level Cognitive Maps for Supporting Between-Floor Spatial Behavior in Complex Indoor Environments

Hengshan Li
University of Maine, hengshan.li@gmail.com

EVALUATION OF MULTI-LEVEL COGNITIVE MAPS FOR SUPPORTING BETWEEN-FLOOR SPATIAL BEHAVIOR IN COMPLEX INDOOR ENVIRONMENTS

By

Hengshan Li

B.S. Wuhan University, 2003

M.S. Wuhan University, 2006

A DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Spatial Information Science and Engineering)

The Graduate School

The University of Maine

May 2016

Advisory Committee:

Nicholas A. Giudice, Associate Professor in Spatial Informatics, School of Computing and Information Science, Advisor

M. Kate Beard-Tisdale, Professor in Spatial Informatics, School of Computing and Information Science

Max J. Egenhofer, Professor in Spatial Informatics, School of Computing and Information Science

Shawn W. Ell, Associate Professor in Psychology

Reinhard Moratz, Associate Professor in Spatial Informatics, School of Computing and Information Science

DISSERTATION ACCEPTANCE STATEMENT

On behalf of the Graduate Committee for Hengshan Li, I affirm that this manuscript is the final and accepted dissertation. Signatures of all committee members are on file with the Graduate School at the University of Maine, 42 Stodder Hall, Orono, Maine.

Signature: ___________________________________ Date: ________________

Dr. Nicholas A. Giudice, Associate Professor in School

of Computing and Information Sciences

LIBRARY RIGHTS STATEMENT

EVALUATION OF MULTI-LEVEL COGNITIVE MAPS FOR SUPPORTING BETWEEN-FLOOR SPATIAL BEHAVIOR IN COMPLEX INDOOR ENVIRONMENTS

By Hengshan Li

Dissertation Advisor: Dr. Nicholas A. Giudice

An Abstract of the Dissertation Presented
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Spatial Information Science and Engineering)
May 2016

People often become disoriented when navigating in complex, multi-level buildings. To efficiently find destinations located on different floors, navigators must refer to a globally coherent mental representation of the multi-level environment, which is termed a multi-level cognitive map. However, there is a surprising dearth of research into underlying theories of why integrating multi-level spatial knowledge into a multi-level cognitive map is so challenging and error-prone for humans. This overarching problem is the core motivation of this dissertation.

We address this vexing problem in a two-pronged approach combining study of both basic and applied research questions. Of theoretical interest, we investigate questions about how multi-level built environments are learned and structured in memory. The concept of multi-level cognitive maps and a framework of multi-level cognitive map

development are provided. We then conducted a set of empirical experiments to evaluate the effects of several environmental factors on users' development of multi-level cognitive maps. The findings of these studies provide important design guidelines that can be used by architects and help to better understand the research question of why people get lost in buildings. Related to application, we investigate questions about how to design user-friendly visualization interfaces that augment users' capability to form multi-level cognitive maps. An important finding of this dissertation is that increasing visual access with an X-ray-like visualization interface is effective for overcoming the disadvantage of limited visual access in built environments and assists the development of multi-level cognitive maps. These findings provide important human-computer interaction (HCI) guidelines for visualization techniques to be used in future indoor navigation systems.

In sum, this dissertation adopts an interdisciplinary approach, combining theories from the fields of spatial cognition, information visualization, and HCI, addressing a long-standing and ubiquitous problem faced by anyone who navigates indoors: why do people get lost inside multi-level buildings. Results provide both theoretical and applied levels of knowledge generation and explanation, as well as contribute to the growing field of real-time indoor navigation systems.

ACKNOWLEDGEMENTS

I would like to offer my sincere thanks to my advisor Dr. Nicholas Giudice for introducing to me the world of Spatial Cognition. His encouragement and interdisciplinary attitude to the field of spatial cognition smoothed my way from a GIS software engineer to a cognitive scientist and made this multidisciplinary thesis possible. His devoted comments on my work motivated me to think critically and offered significant insights to my research. His humor and kindness helped me conquer the emotional storms and stress during my doctoral research. My heartfelt appreciation also goes out to other members on my dissertation committee, Dr. Kate Beard, Dr. Max Egenhofer, Dr. Shawn Ell, and Dr. Reinhard Moratz. Thanks for the advice of cutting half of the hypotheses mentioned in my proposal so that I could focus more thoroughly on the core research questions of interest. I would like to express my gratitude to Dr. Michael Worboys for his encouragement and valuable advice on my research while I took his courses.

I am particularly grateful to Dr. Richard Corey and Dr. Alexander Salveson Nossum for their discussions and comments on my research. My gratitude extends to all fellow graduates in the School of Computing and Information Science, especially to

Christopher Bennett, Matt Dube, Jake Emerson, Shreyans Jain, Saranya Kesavan, Qinghan Liang, Pathum Mudannayake, Hari Palani, RJ Perry, Monoj Kumar Raja, Shirly Stephen, Balaji Venkatesan, Lisa Walton, Fuyu Xu, and Liping Yang for useful discussions on my research.

I would like to thank all the members in the Virtual Environment and Multimodal Interaction lab (VEMI lab), especially Emily Blackwood, Kendra Bird, Jonathan Cole, Sam Gates, Toni Kaplan, Tim McGrath, Brenden Peters, Ethan Porter, Scott Richards, and John San Diego as well as input from Bill Whalen for useful discussions on my research and for the good times we shared in different meetings in Carnegie. I would also like to express my gratitude for the use of the facilities of the lab and thanks to all for taking part in my experiments.

I would like to thank my English tutor Ms. Samantha Le for helping me to improve my English. I also thank my former advisors Dr. Xinyan Zhu and Dr. Hong Fan who supported me to study abroad. I particularly want to thank my wife, Nan Li, and my parents, Jianguo Li and Yunai Li, for their unconditional love and support.

This research was supported by the National Science Foundation under Grant number IIS-0916219, the University of Maine Janet Waldron Doctoral Research Fellowship (UMDRF), and the University of Maine Chase Distinguished Research Assistantship (CDRA). I am grateful for the funding I received.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

Public buildings such as hospitals, libraries, shopping malls, airports, parking facilities, etc., are becoming more complex with many aboveground floors and underground levels. As a case in point, the growing size of malls makes these structures seem like an ‘indoor city’, meaning that they are large and cognitively complex environments with many possible destinations and heavy pedestrian traffic (Uzzell, 1995). Multi-level buildings have the advantage of more efficient use of land space, particularly where space is limited or expensive, and are less expensive to cool or heat compared to a more sprawling single-level structure. However, these complex multi-level buildings often cause navigators to become frustrated, disoriented, or lost during navigation, especially when traversing between floors. It is widely accepted that to efficiently reach a destination in complex environments without becoming lost, navigators rely on the support of cognitive maps—an enduring, observer-free spatial representation of the environment (O’Keefe & Nadel, 1978; Tolman, 1948). Similarly, to accurately and efficiently find targets located on different floors, people must form a globally coherent mental representation of the multi-level built environment, termed a *multi-level cognitive* *map*. Previous literature on navigation in multi-level built environments has shown that the integration of multi-level spatial knowledge acquired from different floors into a multi-level cognitive map is a challenging spatio-cognitive task for human spatial cognition. For instance, navigators have been shown to be significantly less accurate when pointing to locations between floors than within a single floor, and inter-floor knowledge has been argued as the cause of disorientation in both physical and virtual environments (Carlson, Hölscher, Shipley, & Dalton, 2010; Giudice & Li, 2012; Hölscher, Meilinger, Vrachliotis, Brösamle, & Knauff, 2006; Li & Giudice, 2013; Montello & Pick, 1993; Passini, 1992; Richardson, Montello, & Hegarty, 1999; Soeda, Kushiyama, & Ohno, 1997). Given the aforementioned literature highlighting the challenges of between-floor pointing and navigation, there is a surprising dearth of research into underlying theories of why integrating multi-level spatial knowledge into a multi-level cognitive map is so challenging and error-prone for humans. This overarching problem is the core motivation of this thesis.

A second core motivation of the thesis is to address solutions for the indoor navigation problem by developing and evaluating user-friendly visualization interfaces to facilitate cross-level spatial behaviors such as between-floor pointing and wayfinding. Wayfinding is the process of determining and following a path between an origin and a destination through the environment (Golledge, 1999). Many users experience challenges wayfinding accurately and efficiently without effective navigation systems in complex buildings. “The hope is that cognitive science will lead to a better understanding of the human mind, of teaching and learning, of mental abilities and of the development of intelligent devices that can augment human capabilities in important and constructive ways” (Norman, 1981). There is, therefore, a significant need for us to empirically study the very real issue that people get lost inside buildings and to develop perceptually-salient and user-friendly interfaces for supporting human wayfinding as well as other cross-level spatial behaviors in multi-level built environments.

1.2 Goals, Research Questions, Hypotheses, and Evaluations

The primary goal of this thesis is to elucidate empirical evidence that provides valuable insights for addressing the question of why integrating multi-level spatial knowledge into a multi-level cognitive map is so challenging and error-prone for humans. To this end, we defined the concept of multi-level cognitive maps and developed a theoretical framework of multi-level cognitive map development to interpret relevant psychological mechanisms, processes, and knowledge structures involved in learning a new multi-level built environment. This thesis has the following two key goals:

(1) Improving our understanding of how environmental factors affect human mental representation of multi-level built environments (spatial cognition aspect). The

outcomes of the proposed experiments relating to this goal will provide evidence-based design interventions for architects.

(2) Developing and evaluating user-friendly visualization interfaces that augment humans' ability to build multi-level cognitive maps (visualization techniques aspect). The findings of the proposed experiments will provide new human-computer interaction (HCI) principles for cognitively motivated visualization techniques for developing indoor navigation systems.

The first goal provides foundations for the rationale guiding the design of the visualization interfaces, which is the second goal of this thesis. On the other hand, the findings of the second goal validate the hypotheses proposed in the first goal. The two goals lead to the following three key research questions.

(1) How do structural and topological properties of multi-level built environments impact users' development of multi-level cognitive maps?

(2) How do global landmarks (i.e., salient environmental features visible from multiple locations/levels of a building) affect users' mental representation of multi-level built environments?

(3) How do the proposed visualization interfaces assist users' ability to build multi-level cognitive maps?

These three research questions correspond to the following five hypotheses, the background rationale of which will be provided in Chapter 2.

Hypothesis 1: Complex between-floor structural and topological properties impair users' development of multi-level cognitive maps.

Hypothesis 2: Visual access to a global landmark from within a building promotes users' development of multi-level cognitive maps.

Hypothesis 3: If a global landmark is visible from both indoor and outdoor spaces, it will facilitate both cross-level spatial knowledge integration and the integration of outdoor and indoor spaces (OI-spaces).

Hypothesis 4: Using augmented reality (AR) technology to increase visual access to global landmarks will facilitate accurate development of multi-level cognitive maps.

Hypothesis 5: Schematic maps that effectively convey multi-level building information (e.g., providing users with access to between-floor alignment) alleviate the challenge of integrating cross-level spatial knowledge and facilitates accurate development of multi-level cognitive maps.

These hypotheses are investigated throughout this dissertation in seven behavioral experiments (see Table 1.1).

Table 1.1. Five hypotheses are evaluated through Experiments 1-7.

Hypotheses	Experiments
Hypothesis 1	Experiments 1-3
Hypothesis 2	Experiment 4
Hypothesis 3	Experiment 4
Hypothesis 4	Experiments 5-6
Hypothesis 5	Experiment 7

1.3 Scope of Thesis

This thesis research adopts an interdisciplinary approach, combining theories from the fields of cognitive psychology, information visualization, and HCI. The research areas of neuroscience and computational modelling contribute to the findings of this research as well. However, this research is principally empirical and behavioral.

This dissertation is primarily concerned with psychological space—how multi-level built environments are mentally represented and processed. Psychological space includes "concepts which the mind constructs on the basis of reflections on experience, and which would not exist if minds did not exist" (O'Keefe & Nadel, 1978). The focus on mental representation of multi-level built environments distinguishes this thesis research from (1) the research on single-floor indoor wayfinding that deals with spatial cognition issues on a single plane, and (2) the research on indoor navigation that emphasizes modeling of building structures (e.g., connectivity of floor layouts) to support automatic generation of navigation routes (Yang, 2015). This thesis is primarily concerned with how humans learn and represent a multi-level building. Accordingly, the visualization interfaces proposed in this thesis aim to augment users' ability to learn a new multi-level building rather than providing a "crutch" that may actually impair users' spatial abilities in the long-run. In addition, one research endeavor addressed by this thesis is to evaluate new visualization techniques of multi-level buildings for developing indoor navigation systems, which distinguishes this research work from the research on visualization of single-plane outdoor space or a single-floor building.

It is worth noting that although the concepts of multi-level built environments and multi-level indoor environments are interchangeable for the most part in this thesis, multi-level built environments are not necessarily "indoor". For instance, overpasses or flyovers in the outdoor space are also multi-level built environments. This thesis focuses on the "multi-level" property rather than the "indoor" property of multi-level built environments. However, natural multi-layered environments (e.g., trees, burrow systems, caves) are beyond the scope of this thesis.

1.4 Outline of the Thesis

The remainder of this dissertation is organized into the following chapters.

Chapter 2 first reviews existing literature related to the neural representation of three-dimensional space and human wayfinding in multi-level built environments. On this basis, we define the concept of multi-level cognitive maps and propose the framework of multi-level cognitive map development. This is followed by a formal definition of multi-level spatial knowledge (i.e., the knowledge representation of multi-level cognitive maps). Finally, a series of cross-level spatial tasks are designed to evaluate users' formed multi-level cognitive maps.

Chapter 3 first describes the general methods pertaining to the majority of the experiments elaborated in this thesis. Next, we describe two preliminary experiments that investigate the efficacy of using virtual environments to study human mental representation in multi-level built environments. The two studies play an important role in setting the stage of the technical requirements needed in the subsequent work of this thesis.

Chapter 4 describes four experiments (Experiments 1-4) that evaluate the environmental factors, including multi-level structural and topological properties and global landmarks, on the development of multi-level cognitive maps. Finally, the implications of the findings for architectural design are provided.

Chapter 5 describes three experiments (Experiments 5-7) that evaluate several visualization interfaces on the development of multi-level cognitive maps. The implications of the findings for the development of visualization techniques used in indoor navigation systems are provided.

Chapter 6 summarizes the major contributions of this thesis and provides brief sketches of future work.

Finally, the appendix provides a glossary of all concepts used in this thesis.

CHAPTER 2

BACKGROUND

2.1 Multi-level Cognitive Maps

In his seminal book, *The Image of the City* (1960), Kevin Lynch states that "there is a consistent use and organization of definite sensory cues from the external environment. This organization is fundamental to the efficiency and to the very survival of free-moving life" (Lynch, 1960). This mental organization is often modeled as a *cognitive map*, first coined by Tolman (1948), referring to an enduring, observer-free spatial representation of the environment. Tolman used the term cognitive map metaphorically to suggest that animals could use spatial information as if this information was recorded in a map-like manner (Golledge, 1999). In the field of human spatial cognition, it is also widely accepted that to undertake spatial behaviors such as driving to work, or to make spatial decisions like planning a route from home to a shopping mall, we rely on the support of cognitive maps (see reviews in Golledge, 1999; Kitchin & Blades, 2002; Lynch, 1960; Montello & Freundschuh, 2005).

Humans have a long history of curiosity about the nature of space as well as the mental representation of space in the brain. For instance, Newton argued that space is an objective and absolute entity (Newton, 1687), whereas Kant believed that space is an a priori mental framework that our mind uses to coordinate external sensations (Kant, Müller, & Noiré, 1881). However, very little progress had been made in understanding how humans encode physical space in the brain until O'Keefe and Dostrovsky (1971) found that the neurons in the hippocampus, termed "place cells", become active when the animal traverses through a particular location in space (O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1978). This discovery catapulted the hippocampus to the forefront of research on the neural mechanisms of spatial learning and memory. Since then, more than four decades of extensive research in neurophysiology have confirmed that humans have specific kinds of neurons (e.g., place cells in the hippocampus, grid cells, head direction cells, and border cells in the medial entorhinal cortex and adjacent regions) that provide the basis for representation of the spatial environment in the brain (Doeller, Barry, & Burgess, 2010; Hafting, Fyhn, Molden, Moser, & Moser, 2005; Moser, Kropff, & Moser, 2008; Nadel, 1991; O'Keefe, Burgess, Donnett, Jeffery, & Maguire, 1998; O'Keefe & Nadel, 1978). Place cells were suggested to provide animals a dynamic and continuously updated representation of allocentric space and their own positions in that space (Moser et al., 2008; Nadel, 1991; O'Keefe & Nadel, 1978). Grid cells in the medial entorhinal cortex collectively signaled the rat's changing position with a precision similar to that of place cells in the hippocampus, except that each cell had multiple firing fields (Fyhn, Molden, Witter, Moser, & Moser, 2004); grid cells were suggested to possibly be the elements of a metric system for spatial navigation (Hafting et al., 2005).

In recent years, the neural representation of three-dimensional space has attracted increasing attention. For instance, Hayman, Verriotis, Jovalekic, Fenton, and Jeffery (2011) trained rats to walk on a vertical climbing wall ('pegboard') and along a helical track. The experimenters recorded rats' neural activities of place cells and grid cells while they moved on these two setups. The results showed that the firing fields of place cells and grid cells were all elongated in the vertical dimension more than in the horizontal dimensions. Thus, the authors argued that the representation of three-dimensional space in the rat's brain is less precise in the vertical dimension than in the horizontal dimension (Hayman et al., 2011). However, Ulanovsky (2011) counter-argued that anisotropic encoding of three-dimensional space found in this study was due to the rat's body being positioned mostly horizontally and thus their movement patterns were biased toward the horizontal plane. In support, Yartsev and Ulanovsky (2013) found evidence that the hippocampus can represent three-dimensional space by a uniform and nearly isotropic rate code along three axes, as with Egyptian fruit bats (Yartsev & Ulanovsky, 2013). However, no evidence for such 3D representations has been observed in humans. By contrast, Jeffery, Jovalekic, Verriotis, and Hayman (2013) suggested a bicoded representational structure in which space in the plane of locomotion is represented differently from space in the orthogonal axis. On this basis, they argued that "the mammalian spatial representation in surface–traveling animals comprises a mosaic of these locally planar bicoded map fragments rather than a fully integrated volumetric map" (Jeffery et al., 2013), as depicted in Figure 2.1.

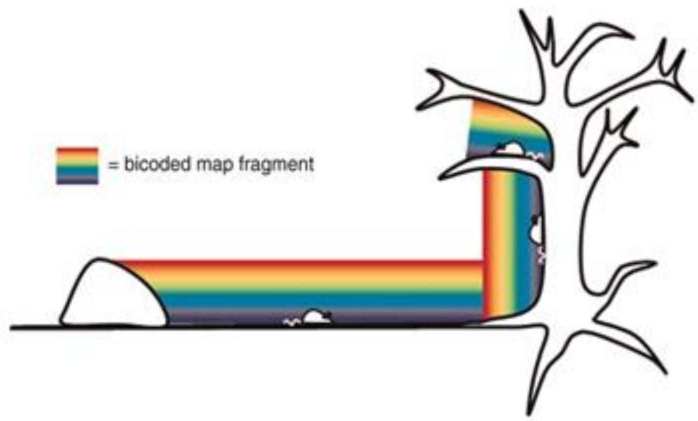


Figure 2.1. Hypothetical structure of the cognitive map in a dimensionally complex environment (adapted from Jeffery et al., 2013).

There has been a lively debate in the literature concerning the efficacy of this bicoded representation, and little is known about whether humans are born with the capacity to construct true 3D spatial representations in the brain (Hölscher, Büchner, & Strube, 2013; Jeffery et al., 2013; Klatzky & Giudice, 2013; Wang & Street, 2013). For instance, Schultheis and Barkowsky (2013) argued that the lack of empirical evidence supporting a three-dimensional volumetric representation in the brain of surface-travelling animals is more indicative of the necessity rather than the ability to maintain such representational structures, as for many spatial tasks the vertical information is either irrelevant to the task or partly redundant with horizontal information. Burt de Perera, Holbrook, Davis, Kacelnik, and Guilford (2013) agreed that the representation of three-dimensional space is probably bicoded. However, they thought it is unlikely that the vertical axis is stored “contextually” without distance or direction metrics. Hölscher, Büchner, and Strube (2013) described that some participants in their experiments consistently reported imagining the indoor environment from an external perspective, with walls and floors like glass, so they argued that the “glass doll house”-like representation requires a “consistent” representation of the vertical dimension. In sum, little hard evidence is available to support whether humans are born with the capacity to construct true three-dimensional (3D) spatial representations in the brain. On the other hand, the bicoded three-dimensional spatial encoding model (Jeffery et al., 2013) is not widely accepted as the form of encoding the three-dimensional world.

This thesis is primarily concerned with human constructed mental representation of multi-level built environments. It does not endeavour to investigate whether humans are capable of constructing a true 3D spatial representation of multi-level built environments, which is beyond the scope of this thesis. Rather, we focus on the functional properties of this mental representation—multi-level cognitive maps should be sufficient for supporting cross-level spatial behaviors in complex multi-level buildings (e.g., accurately pointing and wayfinding to targets located on different floors as discussed in Section 2.4). Thus, in this thesis the concept of multi-level cognitive maps is proposed based on empirical findings and well established conjecture. Previous literature has found clear empirical evidence that humans can encode elevation and a z-axis offset (i.e., vertical distance between floors) in both outdoor and indoor spaces, even if not in a precise 3D manner (Garling, Böök, Lindberg, & Arce, 1990; Tlauka, Wilson, Adams, Souter, & Young, 2007). For example, clear evidence has been found that differences in elevation of terrain are encoded in cognitive maps of outdoor environments (Garling et al., 1990). With regard to multi-level built environments, a growing body of evidence suggests that the integration of multi-level spatial knowledge (learned from different floors) can also be consolidated into a globally coherent mental representation, although this process is found to be challenging and error-prone for humans (Giudice & Li, 2012; Hölscher et al., 2006; Li & Giudice, 2013; Soeda et al., 1997; Tlauka et al., 2007; Vidal, Amorim, & Berthoz, 2004). For instance, Vidal, Amorim, and Berthoz (2004) suggested that human mental representations of buildings could be conceptualized as a set of vertically superimposed 2D cognitive maps having the vertical segments encoded as junctions between the individual maps. Likewise, the results of a series of empirical experiments conducted in large buildings by Hölscher and colleagues suggested that humans have a tendency to memorize multi-level indoor environments as a collection of floors instead of a volumetric representation (Hölscher, Büchner, Meilinger, & Strube, 2009; Hölscher et al., 2006).

On the basis of this empirical evidence, we propose that multi-level cognitive maps have a multi-layered structure, which is different from the concept of a true 3D spatial representation, as the vertical axis of a multi-level cognitive map is not encoded with the same representational structure and fidelity as the horizontal plane. Thus, in this thesis, we partially adopt the bicoded three-dimensional spatial encoding model (Jeffery et al., 2013). However, the bicoded spatial encoding model did not consider how navigators move between floors and how they integrate cross-floor spatial knowledge during vertical travel, both of which are described in this thesis.

In Section 2.2, we propose a framework of multi-level cognitive map development. On this basis, in Section 2.3 we formally define a multi-level cognitive map consisting of (1) a set of super-imposed 2D cognitive maps; (2) between-floor connectivity information (e.g., elevators, staircases, escalators); (3) between-floor alignment information (e.g., indicating what is directly above/below one's current location); and (4) encoding of the z-axis offset (e.g., rough estimates of floor heights). In Section 2.4, we describe our design of a series of cross-level spatial tasks to evaluate users' development of multi-level cognitive maps and illustrate the logical relation between multi-level cognitive maps and cross-level spatial tasks.

2.2 The Framework of Multi-level Cognitive Map Development

To build a cognitive map of a new or unexperienced environment, people usually have to learn it by perceiving surrounding space and thus acquiring spatial knowledge about the environment (Downs & Stea, 1973). Siegel & White (1975) described a theoretical framework for spatial knowledge acquisition over time when learning new environments. In their framework, spatial knowledge is classified into three levels: landmarks, route knowledge and survey knowledge. *Landmarks* refer to distinctive objects or scenes stored in memory. Distant landmarks such as towers or mountain peaks that are visible from a large area of the environment are usually termed *global landmarks*; by contrast, *local landmarks* are visible only from a small distance (Steck & Mallot, 2000). *Route knowledge* is the knowledge of travel paths that connect landmarks. A *route* is "a trace in the environment of a traveled sequence of path segments and turn angles that are followed in order to get from an origin to a destination" (Golledge, 1999). *Survey knowledge* is defined as a configurational representation of spatial relationships between non-linearly-aligned sets of environmental features such as routes and landmarks, organized within a common *spatial reference frame* or *spatial reference system* (Montello, 1998). A related term, *region*, represents perceived and encoded representations in spatial memory in which locations are grouped within a common spatial reference frame (Wiener & Mallot, 2003; Wiener, Schnee, & Mallot, 2005). An important characteristic of survey knowledge is that navigators can infer the spatial relations between locations even if they have never traveled between these locations (Montello, 1998). To acquire survey knowledge of a new environment through active exploration, navigators have to integrate the linearized geometry of the route, landmarks on and off the route, and specific regions through which routes pass, into a configurational layout (Golledge, 1999). The traditional framework of spatial knowledge acquisition argues that (1) landmarks are the first to be acquired, and (2) route knowledge does not contain metric information, e.g., distances and directions. However, Montello (1998) counter-argued that there is no stage in which only landmark or only route knowledge exists, but some configurational knowledge begins to be acquired upon first exposure, which becomes more complete and detailed with increasing experience in the environment.

As discussed above, survey knowledge is based on the integration of acquired spatial knowledge into a common spatial reference system. McNamara, Sluzenski, and Rump (2008) define a *spatial reference system* as a relational system that consists of reference objects, located objects, and the spatial relations that may exist among them. Previous literature generally distinguishes spatial reference systems as either being egocentric or allocentric (Hart & Moore, 1973). With the egocentric reference system, the objects and spatial relationships of the environment are organized with respect to the observer's positions and orientations, whereas in the allocentric reference system, the location and orientation of objects are specified with respect to the environment. Route knowledge is usually obtained from direct experience during navigation from a navigator's perspective, so it is often considered egocentric. Survey knowledge is allocentric, as it is not directly available from first-person perceptual experience and needs to be constructed through the integration of spatial knowledge learned from different locations. As discussed in Section 2.1, cognitive maps refer to observer-free, allocentric spatial representations of an environment, so in this thesis the concepts of users' formed cognitive maps and users' learned survey knowledge are interchangeable for the most part.

With respect to spatial reference systems of multi-level built environments, previous literature has suggested that the intrinsic structural characteristics of multi-level indoor spaces influence one's spatial representation of a building in that they may be coherent locally but not globally (Carlson et al., 2010). To better understand the meaning of "local" and "global" with respect to multi-level built environments, it is necessary to introduce the concept of spatial scale. Montello (1993) identified four scales of space and differentiated between *figural space* (object-sized spaces perceived from one vantage point), *vista space* (room-sized spaces perceived from one vantage point but allowing for head rotation), *environmental space* (perceived by moving through the space) and *geographical space* (experienced from symbolic representations, such as maps).

According to Montello (1993), a room is a *vista space*, as one sees the entire spatial extent from a single vantage point with head rotation. It is similar for lobbies, atriums, etc., but not for most building floors since a floor often has many occlusions that block a navigator's view and limit perceptual access from one point, even if one rotates in place. Thus, most floors of buildings, unless they are completely open, have to be perceived by moving through the space and thus, by definition, are *environmental spaces*. Depending on the spatial scale, spatial reference systems can be divided into two categories: *global reference systems* and *local reference systems* (Gärling, Lindberg, Carreiras, & Book, 1986; Meilinger, Riecke, & Bülthoff, 2013). A *global reference system* is encoded in the entire environmental space, so places and spatial relations can be learned relative to this global reference. By contrast, a local reference system is encoded for a local vista space and it varies from one vista space to another (Gärling et al., 1986). However, local reference systems may be interrelated and serve as elements in higher-level reference systems (McNamara et al., 2008; Meilinger et al., 2013). For instance, the reference systems of rooms may serve as elements in a higher-order reference system defining the spatial relations within a floor while the reference systems of floors subsequently serve as elements in a higher-order reference system within a whole building.

The concept of spatial reference systems is very important to this thesis, as we propose that multi-level cognitive maps are constructed and integrated from local spatial knowledge learned from different floors, and one major research endeavor addressed by this thesis is to investigate how people integrate spatial knowledge learned from different floors (local spatial reference systems) into a multi-level cognitive map (global spatial reference system). Meilinger (2008) proposed the network of reference frames theory to describe the process of integrating local spatial reference frames into a global spatial reference frame in single-plane outdoor space. The basic unit in the network is the allocentric spatial reference frame of a vista space. An edge in the network defines the so-called *perspective shift* that is necessary to move from one reference frame to the next, which consists of both a translation and a rotation component (Meilinger, 2008). In this thesis, we adopt the network of reference frames theory and extend it into multi-level built environments. First, a node in the network of reference frames represents the allocentric spatial reference frame of a floor. Second, the edge in the network (perspective shift) represents a navigator's movement from one floor's reference frame to the next floor. The rotation component of this perspective shift is denoted by γ, as seen in Figure 2.2.

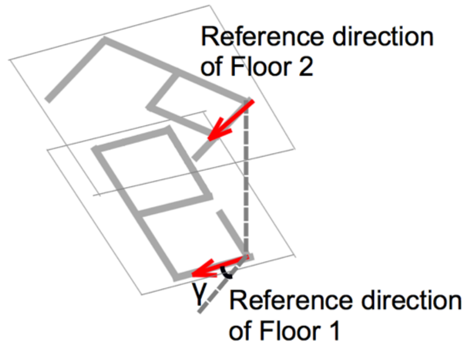


Figure 2.2. Between-floor perspective shift γ. The red arrows indicate the reference directions of the two floors.

Previous literature has discussed three strategies for acquiring survey knowledge of a new environment: (1) active exploration of the environment according to specific rules or using controlled navigational practices such as path integration; (2) using configurations of landmarks to determine navigators' locations and directions; and (3) using secondary information sources such as maps, videos, and verbal descriptions to learn about the environment (Gallistel, 1990; Golledge, 1999; Loomis, Klatzky, Golledge, & Philbeck, 1999). In this thesis, we argue that in order to learn a new multi-level built environment, navigators have to depend jointly on these three strategies. In the following three subsections, we will discuss how navigators integrate cross-level spatial knowledge into a multi-level cognitive map based on the three strategies.

2.2.1 Integrating Cross-level Spatial Knowledge Depending on Path Integration

Learning an environment by experiential procedures such as learning a route and being aware of proximal and distant environmental features is perhaps the most common approach used by humans (MacEachren, 1992). This section discusses how navigators integrate cross-level spatial knowledge through active search and exploration of a multi-level building. We will begin by introducing a few important terms associated with navigation, illustrated in Figure 2.3.

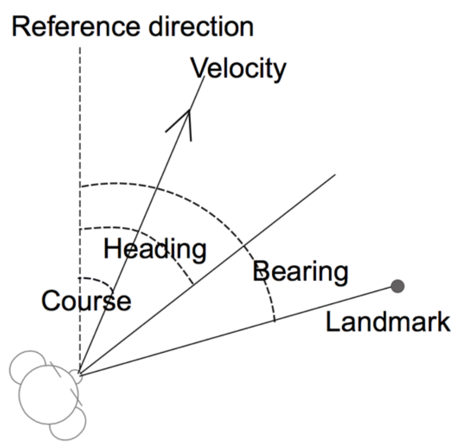


Figure 2.3. Depiction of navigational terms (adapted from Loomis, Klatzky, Golledge, & Philbeck, 1999).

As shown in Figure 2.3, *reference direction* refers to the orientation of a reference frame. In outdoor space, navigators usually use north as the reference direction based on azimuthal cues such as the sun. In this thesis, we focus on multi-level built environments, so a reference direction refers to the orientation of a local reference frame, perhaps a room or a floor. Previous literature has suggested that the orientation of a local reference frame usually originates from the initial experience with that space (e.g., after entering a floor), the main experienced orientation within this space, and its overall structure (e.g., the main orientation of a room or a floor) (Kelly & Mcnamara, 2008; McNamara et al., 2008; Meilinger et al., 2013; O'Keefe, 1991). *Course* and *heading* are the direction of a navigator's velocity vector and facing direction, respectively, measured with respect to the reference direction, and *bearing* refers to the direction from the navigator to a landmark (Gallistel, 1990; Loomis, et al., 1999).

With respect to multi-level buildings, a prominent and fundamental characteristic of learning a multi-level built environment is considering the *vertical transitions*—navigators use *vertical connectors* (e.g., elevators, staircases, escalators) to navigate between floors. It is necessary to define a few concepts involved in this process that are important for illustrating how navigators integrate between-floor spatial knowledge.

The *transition point* is defined in this thesis as representing a point where navigators enter or exit a floor. The notion of transition point is different from the related term, decision point, which usually refers to the point where two route-segments meet or the intersection of two or more corridors or travel paths (Richter & Klippel, 2005). Transition points are the connecting points of two floors. An outdoor transition point is the intersection between two adjacent regions' common boundary and a route that goes through the two regions, whereas a transition point in multi-level built environments is the point where users pass through a *between-floor portal* such as an elevator door to enter or exit a floor by a vertical connector. A transition point has one direction based on the orientation of a between-floor portal. Given that we focus on between-floor spatial knowledge integration, and for the sake of simplicity, we assume that a navigator's velocity vector and facing direction are the same as the transition point's direction when entering or exiting a between-floor portal, as pictured in Figure 2.4.

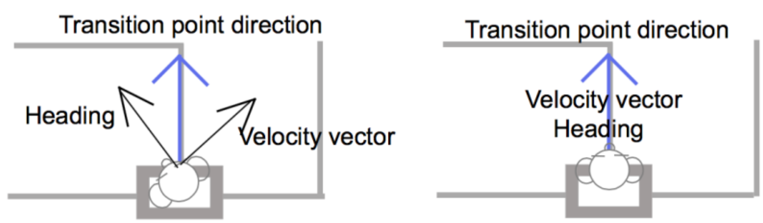


Figure 2.4. A navigator's velocity vector, facing direction and transition point direction when entering or exiting a between-floor portal.

When people navigate between floors, they will pass a pair of transition points. For example, as shown in Figure 2.5, there is one pair of transition points ($p1$ and $p2$) connecting Floor 1 and Floor 2. Navigation between the two floors involves a vertical transition offset (h), a horizontal transition offset (d), and a transition angular offset (α).

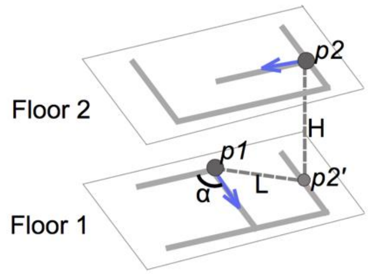


Figure 2.5. Transition points in multi-level built environments. The blue arrows indicate the directions of a pair of transition points (*p*1 and *p*2).

The *vertical transition offset* is the z-axis offset between this pair of transition points located on different floors. The *horizontal transition offset* is the offset between the transition point (*p1*) and the projection of the corresponding transition point (*p2'*) on the former transition point's floor (e.g., floor 1). If the two transition points are vertically aligned, as when an elevator connects the pair, the horizontal offset L is 0. The *transition angular offset* is the difference between navigators' facing direction at a pair of transition points. The transition angular offset is termed the *between-floor heading shift* (denoted by α) in this thesis. In addition, the term *portal-floor heading shift* is used in this thesis to represent the angular offset between the reference direction of a floor and a navigator's heading when entering or exiting a between-floor portal. The portal-floor heading shifts on two floors are denoted as $\beta 1$ and $\beta 2$ respectively, as shown in Figure 2.6.

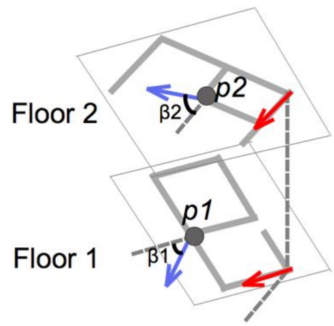


Figure 2.6. Portal-floor heading shifts (*β1* and *β2*) on two floors. The red arrows indicate the reference directions of the two floors. The blue arrows indicate the directions of a pair of transition points.

So far, we have introduced four heading shifts during vertical transition: (1) a perspective shift *γ*; (2) a between-floor heading shift α; and (3) two portal-floor heading shifts (*β1*, *β2*). To maintain orientation during vertical transitions, navigators have to update heading, which can be accomplished either by direct sensing or by integrating turn rate (Loomis, et al., 1999). However, navigators usually do not have direct perception of both floors during vertical transition due to occlusion from elevator shafts or stairwells. Instead, they have to sense rotary accelerations, depending on proprioceptive, vestibular, or optic cues, and doubly integrate this information to obtain rotational displacements (Loomis, et al., 1999). This process is termed *path integration* (also called *dead-reckoning*), referring to the updating of position and heading on the basis of velocity and acceleration information (Loomis, et al., 1999; Loomis et al., 1993b).

It is worth noting that during vertical transition navigators still have direct perception through visual and other sensing of their immediate surroundings such as the space within an elevator. The perceived information provides important cues about the between-floor heading shift. For instance, if an elevator has separate entrance and exit doors that are offset by 90°, they can directly perceive that the between-floor heading shift is 90°. However, when a multi-level building has no between-floor visual access and no visual access between indoor and outdoor spaces (called OI-spaces), navigators have to depend on the between-floor heading shift and the two portal-floor heading shifts $\beta 1$ and $\beta 2$ in order to integrate cross-level spatial knowledge, as illustrated in Figure 2.7. However, during vertical transition the portal-floor heading shift is not directly perceivable. Thus, we argue that path integration plays an important role in supporting the integration of cross-level spatial knowledge during vertical transition. Likewise, Loomis, et al. (1999) described a few natural occurring examples of using path integration, stating that: "Fire fighters entering a smoke-filled building depend on path integration for positional awareness. Cave explorers and divers face the challenge of having to do path integration in all three spatial dimensions. Finally, pilots flying under instrument conditions while being vectored by ground controllers must engage in imagined path integration to maintain an estimate of their relationship to their destination and any hazardous terrain" (p. 150). In all these scenarios, although people still have direct perception of their immediate surroundings, path integration plays a critical role in positional and directional awareness. Thus, for simplicity in this thesis we use the term path integration to describe the process of integrating cross-level spatial knowledge during vertical travel. The process of calculating the perspective shift γ based on the between-floor heading shift α and two portal-floor heading shifts β1 and β2 is illustrated in Figure 2.7: (1) γ = β2 - (β1 - α), (2) γ = β2 - (β1 + α), (3) γ = β2 - (α - β1), (3) γ = (β1 - α) + β2, (4) γ = (β1 - α) - β2, (5) γ = (β1 + α) - β2, (6) γ = (α - β1) - β2, and (7) γ = β2+ (β1 + α).

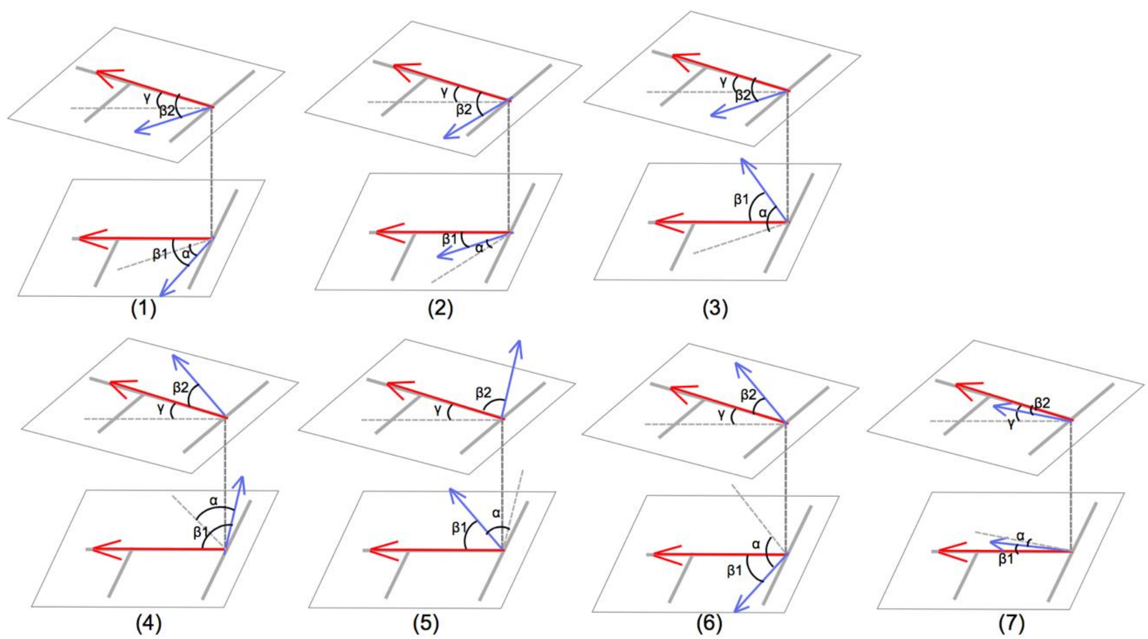


Figure 2.7. Calculating perspective shift γ based on the between-floor transition shift α and two portal-floor angular offsets (*β1* and *β2*).

It is not surprising that confusing staircases have been identified as a main cause for people getting lost inside buildings (Hölscher et al., 2006), as the path integration process that we argue is critical for integrating cross-level spatial knowledge is very challenging (Etienne & Jeffery, 2004; Klatzky, Beall, & Loomis, 1999; Loomis, et al., 1999; Loomis et al., 1993a). After navigators learn the between-floor perspective shift γ, they can calculate between-floor alignment information based on the between-floor perspective shift γ and the horizontal transition shift (*d*) of a pair of transition points (*p1*, *p2*), as seen in Figure 2.8.

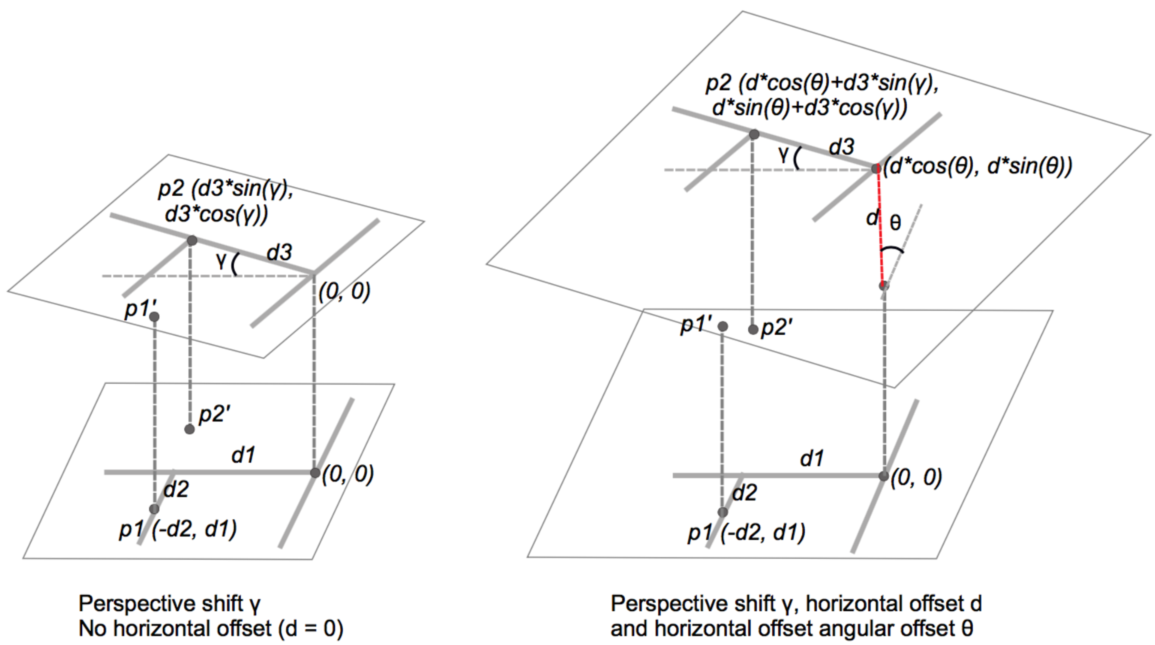


Figure 2.8. Calculating between-floor alignment information based on the perspective shift γ and the horizontal transition offset *d.*

Given that complex multi-level structural and topological properties such as misalignment between floors may increase the difficulty of path integration and subsequent between-floor alignment calculation, in this thesis we propose that complex between-floor structural and topological properties impair users' development of multi-level cognitive maps (*Hypothesis 1*). This hypothesis is evaluated in Chapter 3 through Experiments 1-3.

2.2.2 Integrating Cross-level Spatial Knowledge based on Global Landmarks

In the previous section, we discussed the situation of when a multi-level building has no between-floor visual access and no visual access between OI-spaces. In such cases, navigators have to obtain rotational displacements by updating heading information during vertical transition in order to integrate cross-level spatial knowledge. However, this path integration process has been found to be difficult for human spatial cognition (see review by Loomis, et al., 1999). In this section, we will explore what happens when a multi-level building has visual access to a global landmark such as a nearby prominent building or an atrium as well as how navigators may use the global landmark to integrate cross-level spatial knowledge.

Previous literature has shown that navigators use configurations of landmarks to determine their location or heading in a process called *piloting* (Gallistel, 1990; Loomis, et al., 1999). In his seminal book, *The Organization of Learning* (1990), Gallistel discussed several methods of piloting, such as computing position using bearing and distance to a single landmark, computing position using distances to multiple landmarks (trilateration), and computing position using bearings or bearing differences to multiple landmarks (triangulation). For instance, as illustrated in Figure 2.9, navigators can use the estimated distance (*d*) that the observer has moved between two positions (*p1* and *p2*) and the two bearings of the two points (θ1 and θ2) for triangulation of the distance (D) between *p2* and the landmark.

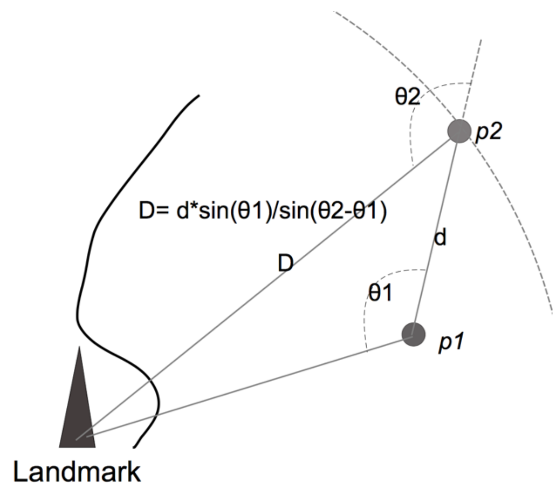


Figure 2.9. Computing distance (D) between *p2* and the landmark using estimated distance (*d*) between two positions (*p1* and *p2*) and two bearings of the two positions (*θ1* and *θ2*; adapted from Gallistel, 1990).

Similarly, in multi-level built environments, navigators can calculate the distance and angle between two positions (*p1* and *p2*) located on two floors based on bearings of the two positions ($\theta 1$ and $\theta 2$) and estimated distances ($d1$ and $d2$) between the two positions and the landmark, as demonstrated in Figure 2.10.

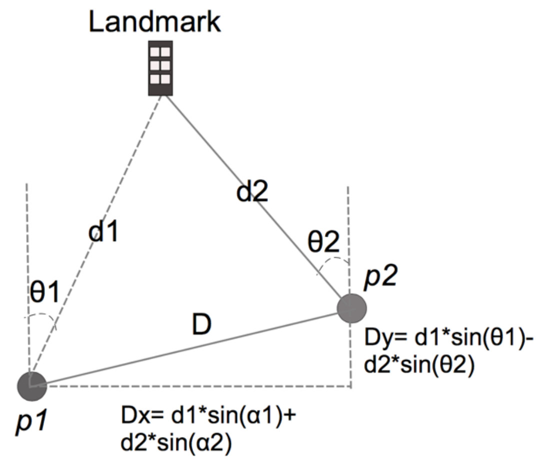


Figure 2.10. Computing distance (D) between two positions ($p1$ and $p2$) based on bearings of the two positions ($\theta 1$ and $\theta 2$) and estimated distances ($d1$ and $d2$).

Although this formal process of bearing-based distance and direction calculation is more accurate and precise than the path integration process introduced in the previous section, the formal process is not intuitive and unlikely to be used in our daily lives. In addition, accurate distance and angle estimation between two positions located on two floors is not critically important or necessary for accurate cross-level spatial knowledge integration, as in this process, navigators are primarily concerned with the between-floor perspective shift γ. Next, we will discuss how navigators use global landmarks to learn between-floor perspective shift γ.

The example depicted in Figure 2.11 assumes that there are two positions (*p1* and *p2*) located on two floors and there is no direct visual access between the two positions. However, if navigators can see a global landmark from both locations, they can calculate the between-floor perspective shift γ based on the two bearings (*θ1* and *θ2*) of the two positions (i.e., angular offsets between the orientation of the global landmark and the two floors' reference directions). For the example shown in Figure 2.11, perspective shift $\gamma = \theta 1 + \theta 2$.

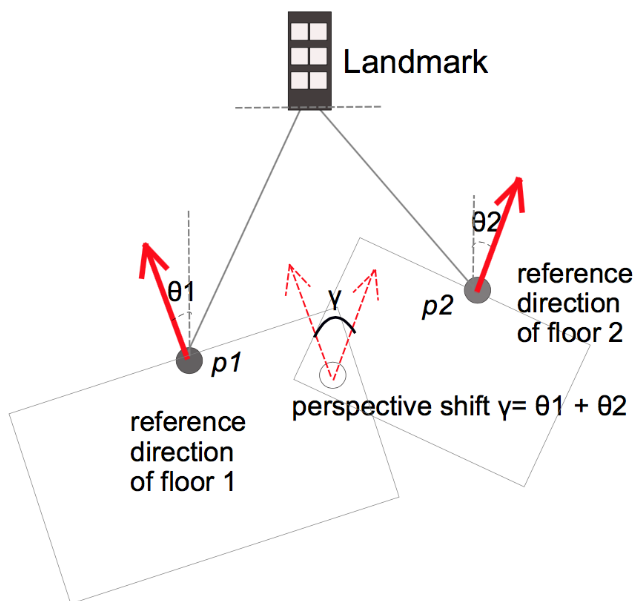


Figure 2.11. Estimate between-floor perspective shift γ based on the bearings of two positions (*p1* and *p2*).

As discussed in the previous section, in order to integrate cross-level spatial knowledge based on path integration, navigators need three angles including the between-floor heading shift *α* and two portal-floor heading shifts *β1* and *β2*. In contrast, navigators need only two angles (two bearings *θ1* and *θ2*) to integrate cross-level spatial knowledge based on global landmarks. Thus, cross-level spatial knowledge integration based on global landmarks requires less computation than that solely based on path integration. In addition, the perceived information (e.g., visual access to a global landmark) could be helpful to confirm or validate the heading updating process and to reduce spatial uncertainty accumulated in the path integration process. Thus, we anticipate that visual access to a global landmark from within a building promotes users' development of multi-level cognitive maps (*Hypothesis 2*). This hypothesis is evaluated in Chapter 4 through Experiments 4-5.

Previous literature on qualitative spatial reasoning (QSR) has discussed both reasoning about orientations in a global reference frame (using cardinal directions such as North) and reasoning about relative directions in local reference frames (using relative directions such as left/right and front/back; see Clementini, Felice, & Hernández, 1997; Freksa, 1992; Goyal & Egenhofer, 1997; Moratz & Ragni, 2008). This reasoning process can be refined based on the knowledge of distance (see Goyal & Egenhofer, 2001; Moratz & Ragni, 2008; Moratz & Wallgrün, 2012). Although the formalization of positional reasoning in multi-level built environments is outside the scope of this thesis, previous research on QSR offers important implications for cross-level spatial knowledge integration: navigators can use global landmarks to estimate relative direction between two positions located on two floors, which is proposed to be cognitively easier than the aforementioned bearing-based direction calculation.

Figure 2.12 shows an example in which there are two positions (*p1* and *p2*) located on two floors and no direct visual access between these two positions. Navigators can see the global landmark, a nearby prominent building, from both positions. If they use the orientation of the building as the local reference direction, they can learn that *p1* is located at the left/back of the landmark and *p2* is located at the right/back of the landmark. According to previous research on reasoning of relative directions, navigators can infer that *p2* is located either at the right, right/front, or right/back of *p1*. In addition, if navigators use estimated distances between the two positions and the global landmark, the estimated relative direction between the two positions will be further refined, providing important cues for cross-level spatial knowledge integration and the integration of OI-spaces. Thus, we anticipate that if a global landmark is visible from both indoor and outdoor spaces, it will facilitate both cross-level spatial knowledge integration and the integration of OI-spaces (*Hypothesis 3*). This hypothesis is evaluated in Chapter 4 using Experiment 5.

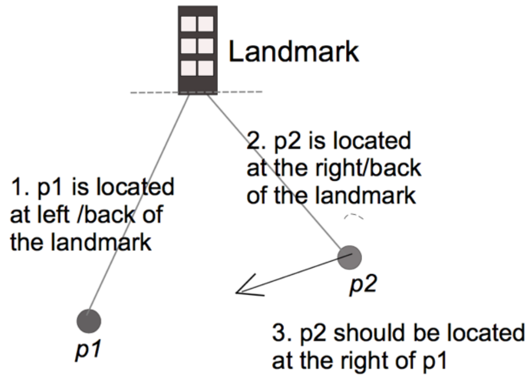


Figure 2.12. Estimating relative direction between two positions (*p1* and *p2*).

2.2.3 Using Visualization Interfaces to Facilitate Cross-level Spatial Knowledge Integration

In the previous section, we described the rationale for our prediction that global landmarks are anticipated to promote users' development of multi-level cognitive maps and to facilitate the integration of OI-spaces. However, global landmarks are often not available in multi-level indoor environments, due to: (1) interior objects such as walls, ceilings, and other obstacles limiting visual access, and (2) the external windows or large atriums that might be used to facilitate access are frequently only visible from specific locations in the building. As a result, the advantage of global landmarks serving as a fixed spatial frame of reference is often greatly reduced when learning and navigating through indoor environments (Giudice, Walton, & Worboys, 2010). If visual access to global landmarks is found to facilitate the development of a multi-level cognitive map, as we predict, the question remains as how to best leverage this benefit for the majority of complex buildings that lack direct visual access to such landmarks. It is obviously impractical to modify the physical building to increase access but an alternative and economical solution is to use visualization techniques such as Augmented Reality (AR). AR technology can be used to superimpose virtual information on the physical environment from a perception-friendly first-person perspective and thus enhance users' spatial awareness of the environment by showing occluded information that they otherwise cannot directly perceive (Dey & Sandor, 2014). If we can use AR technology to increase visual access to global landmarks, the benefit of these cues for providing a fixed frame of reference can be extended to all matter of complex multi-level buildings and thereby facilitate users' development of multi-level cognitive maps (*Hypothesis* 4). This issue is investigated in Chapter 5 through Experiments 5-6.

In addition to AR visualizations, we investigate how *schematic maps* as implemented on mobile devices affect users' development of multi-level cognitive maps. *Schematic maps* refer to maps that omit unnecessary details and use simplified structures to represent meaningful entities and spatial relationships, such as public transportation maps, metro maps, and tourist city maps. These schematic maps are argued to be cognitively efficient for supporting users in learning environments (Avelar & Hurni, 2006; Casakin, Barkowsky, Klippel, & Freksa, 2000; Klippel, Richter, Barkowsky, & Freksa, 2005; Ware, Taylor, Anand, & Thomas, 2006; Schmid, Richter, & Peters, 2010; Schmid, 2008). Most currently available indoor digital maps such as Google Indoor maps, however, can only show one floor at a time, meaning users can obtain only visual access to individual floors and thus have no direct means for integrating knowledge between floors. Although prevalent, this type of map requires that users manually integrate spatial knowledge between multiple layers over time, and this temporal and spatial integration is likely hard, inaccurate, and leads to errors, as are known for the development of multi-level cognitive maps (Carlson et al., 2010; Giudice & Li, 2012; Hölscher et al., 2006; Li & Giudice, 2013; Montello & Pick, 1993; Passini, 1992; Richardson et al., 1999; Soeda et al., 1997). In this thesis, we propose that schematic maps that effectively convey multi-level building information (e.g., providing users with access to between-floor alignment) alleviate the challenge of integrating cross-level spatial knowledge (*Hypothesis 5*). This hypothesis is evaluated in Chapter 5 using Experiment 7.

AR visualization usually provides perception-friendly egocentric spatial information in order to enhance users' spatial awareness of the environment, but it requires navigators to integrate egocentric spatial knowledge obtained from different places into an allocentric cognitive map. Schematic maps directly provide allocentric spatial information regarding the environment, but they are often difficult to learn and can lead to a phenomenon well-known in the spatial cognition literature as *the alignment* *effect*. It occurs when the reference direction (e.g., the upward direction) in a visual or haptic map is misaligned with a navigator's facing direction in the environment, causing the judgment of the direction of environmental features represented on the map to be slower and less accurate than when the map is aligned with their facing direction (Giudice, Betty, & Loomis, 2011; Levine, Jankovic, & Palij, 1982; Mou, McNamara, & Valiquette, 2004; Presson, DeLange, & Hazelrigg, 1989; Waller, Montello, Richardson, & Hegarty, 2002).

Schematic maps usually maintain a global reference frame (e.g., using a north-up display), thus emphasizing global awareness, whereas AR visualizations often employ an egocentric frame, thereby improving local guidance. However, emphasizing solely global awareness or local cues is not optimal for cognitively motivated visualization techniques, as both types of information are important for navigation (Taylor et al., 2008). Therefore, we studied both AR visualizations (Experiments 5-6) and schematic maps (Experiment 7) in aiding the development of multi-level cognitive maps.

The difference between the effects of the two types of visualization interfaces is outside the scope of this thesis. However, it is worth noting that an important design principle of cognitively motivated visualization techniques for indoor navigation systems is to systematically combine the advantages of different types of visual interfaces (e.g., schematic maps and AR visualizations) in what has been called a details-on-demand visual interface in a previous study (Li & Giudice, 2012a; see further discussion in Section 6.2.2).

2.3 Multi-level Spatial Knowledge

The previous sections discussed three strategies of cross-level spatial knowledge integration, including active exploration (path integration; Section 2.2.1), using configurations of landmarks to determine navigators' locations and directions (piloting; Section 2.2.2), and learning about the environment through secondary information sources such as schematic maps and AR visualizations (Section 2.2.3). The final product of this integration process is referred to in this thesis as *multi-level spatial knowledge*. Multi-level spatial knowledge is constructed by integrating single-level spatial knowledge from different floors (local reference systems) into a common spatial reference frame (global reference system) and described in terms of multi-level landmarks, multi-level route knowledge and multi-level survey knowledge.

Multi-level landmarks are distinctive objects or scenes that are visible from multiple locations/levels of a multi-level built environment. In this thesis, multi-level landmarks are visible from all floors of a building and thus are interchangeable with global landmarks hereinafter for consistency. *Multi-level route knowledge* is the knowledge of travel paths that connect between-floor locations. Navigators obtain this multi-level route knowledge from direct experience during navigation, meaning that multi-level route knowledge is based on an egocentric spatial reference frame. By contrast, *multi-level survey knowledge* is a configurational representation of the metric spatial relationship between environmental features across multiple floors organized within a common spatial reference frame. Multi-level survey knowledge is allocentric, as it needs to be constructed through the integration of spatial knowledge learned from different locations/floors. In the following paragraphs, we first define single-level survey knowledge (i.e., survey knowledge formed in a single-plane environment) and then extend the definition to multi-level survey knowledge.

In this thesis, a configurational representation of a region/floor (single-level survey knowledge) is described as consisting of: (1) name of the region (*ID*); (2) a reference direction vector ($\overrightarrow{RD}$); (3) a finite collection of places $P = \{p_1, p_2, \dots, p_n\}, p_i \in \mathbb{R}^2$, where $\mathbb{R}^2$ is a Euclidean space; (4) a finite collection of routes $R = \{r_1, r_2, \dots, r_n\}$, r_i is a route described by an ordered set of places $\{x_1, x_2, \dots, x_n\}, x_i \in P$, such that $(x_1, x_2), (x_2, x_3), \dots, (x_{n-1}, x_n)$ are segments of the route and the x_i are distinct; and (5) metric relations of these places in terms of distances and directions, $D = \{\overrightarrow{p_1 p_2}, \overrightarrow{p_3 p_4}, \dots, \overrightarrow{p_i p_j}\}, p_i, p_j \in P, i \neq j.$

Cognitive maps should be able to support both metric relations and topological relations (Egenhofer & Franzosa, 1991), although the learned metric information is typically fragmented and distorted (McNamara et al., 2008; Sadalla & Montello, 1989; Stevens & Coupe, 1978; Tversky, 1981, 1992). This thesis, however, is primarily concerned with metric relations in terms of distances and directions. As discussed in Section 2.2, if navigators have formed accurate survey knowledge of a region, they can infer the spatial relations between locations even if they have never traveled between these locations before (Montello, 1998), meaning that (1) navigators can accurately point between places within the region (*single-level pointing task*), and (2) navigators can accurately and efficiently perform wayfinding between places within the region using the shortest path (*single-level wayfinding task*). Wayfinding refers to a set of spatial processes allowing for the determination and execution of routes between an origin and a destination without prior knowledge of the route (See Golledge, 1999 for an excellent review of the literature on wayfinding behavior).

On the basis of the definition of single-level survey knowledge, multi-level survey knowledge is then described as consisting of:

1) A set of super-imposed single-level survey knowledge presentation. Let L_i be the survey knowledge of level i in a set of super-imposed single-level survey knowledge $LS = \{L_1, L_2, \dots, L_n\}, n \in N, n \le floor\ \ numbers$. If multiple floors of a building have the same (or similar) layouts, these floors could be mentally represented as one level in the multi-level cognitive map. In this thesis, different floors usually have

distinctive floor layouts, so for simplicity each single-level spatial knowledge representation corresponds with a single building floor.

2) Between-floor connectivity information.

As described in Section 2.2.1, when people navigate between floors, they must pass through a pair of transition points. Transition points are denoted as $T = \{\vec{t}_{1j}, \vec{t}_{2j}, \ldots, \vec{t}_{ij}\}$, where $\vec{t}_{ij}$ is a vector with a starting point $t_{ij} \epsilon\ P_j\ , P_j\ \subset L_j$. The direction of vector $\vec{t}_{ij}$ is the orientation of the corresponding between-floor portal. Between-floor connectivity information is a finite collection of paired transition points. Let $\overrightarrow{t_{ij}}$ be a transition point $\overrightarrow{t_i}$ on floor *j* and $\overrightarrow{t_{mk}}$ be a transition point $\overrightarrow{t_m}$ on floor *k*. Between-floor connectivity information is denoted as C_{jk}.

$C_{jk} = \{(\overrightarrow{t_{11}}, \overrightarrow{t_{12}}), (\overrightarrow{t_{21}}, \overrightarrow{t_{22}}), \ldots, (\overrightarrow{t_{ij}}, \overrightarrow{t_{mk}})\}, i,\ j, k \in N, j \neq k, j, k \leq$ $floor\ numbers,\ t_{ij} \in\ P_j, P_j \subset L_j,\ t_{mk} \in\ P_k, P_k \subset L_k$. The t_{ij} are not distinct, as one transition point can be connected to multiple transition points located on a different floor.

3) Between-floor alignment information.

Let p_{ij} be a place p_i on floor j, and let p'_{ijk} be the vertical projection of place p_{ij} on floor k, where between-floor alignment information is a finite collection of paired places between p_{ij} and p'_{ijk}. Between-floor alignment information is denoted as

$A_{jk}. A_{jk} = \{(p_{11}, p'_{112}), (p_{12}, p'_{123}), \ldots, (p_{ij}, p'_{ijk})\}, i, j, k \in N, j, k \leq floor\ \ numbers, j \neq k, p_{ij} \in \ \ P_j, P_j \subset L_j, p'_{ijk} \in \ \ P_k, P_k \subset L_k.$

4) Encoding of the z-axis.

$H = \ \{h_{12}, h_{21}, h_{26}, h_{64}, \ldots, h_{jk}\}, j, k \in N, j \neq k, and\ \ j, k \leq floor\ \ numbers.$

As described in Section 2.1, there has been a lot of debate regarding the encoding resolution and fidelity of the z-axis in three-dimensional space. The bottom line, however, is that previous literature has shown that humans can roughly estimate distance between floors, although the estimations are distorted, with "relative downward errors in upward judgments and relative upward errors in downward judgments" (Jeffery et al., 2013; Tlauka et al., 2007; Wilson, Foreman, Stanton, & Duffy, 2004). For instance, for a three-story building with each floor being 12 feet high, if navigators stand on floor 1 and make upward distance judgments, their estimates of the floor height of the three-story building might be 11 feet (floor 1), 10 feet (floor 2) and 9 feet (floor 3). However, if navigators stand on floor 3 and make downward judgments, the floor heights of the building might be estimated as being 9 feet (floor 1), 10 feet (floor 2) and 11 feet (floor 3). Investigating the encoding and distortion of the z-axis dimension will be a research topic in a future project but is not the focus here (see Chapter 6). This thesis is primarily concerned with the integration of superimposed single-level spatial knowledge, and we assume that navigators can roughly estimate distance between floors.

In this thesis, multi-level survey knowledge is defined as a finite collection (a set) of places, routes, connectivity, and vertical alignment as well as other metric relations in terms of directions and distances, because "knowledge representations of space are probably not best conceived of as coherent, unchanging wholes, but rather as conglomerations of information drawn from different sources and modalities and pulled together for a particular purpose" (Mark, Freksa, Hirtle, Lloyd, & Tversky, 1999). Although the concepts of multi-level survey knowledge and multi-level cognitive maps are frequently used interchangeably in this thesis, it is worth noting that multi-level cognitive maps usually refer to the neural representation of multi-level built environments, whereas multi-level survey knowledge usually refers to knowledge representation such as positions, directions, and connectivity that is represented and retrieved for a specific spatial task. In the next section, we will introduce a series of cross-level spatial tasks used to evaluate users' formed multi-level survey knowledge.

2.4 Cross-level Spatial Tasks

In this thesis, we designed three types of cross-level spatial tasks to evaluate users' formed multi-level survey knowledge.

Task 1: *Between-floor pointing task.* Users point from position *p1* on one floor to position *p2* on another floor. In the experiments described here, users were required to point horizontally by imagining that the two positions were on the same plane. Note that if users can point horizontally to the target located on a different floor, they can also point directly to the target in the three-dimensional space, as they can roughly estimate distance between floors.

Task 2: *Between-floor wayfinding task.* Users navigate from position *p1* on one floor to position *p2* on another floor using the shortest path.

Task 3: *Vertical alignment tasks.* There are three types of vertical alignment tasks investigated by the experiments in this thesis: (1) *Drilling task*: users indicate which object or landmark (e.g., room) is directly above/below their current location (Experiments 4-6), (2) *vertical navigation task*: users navigate from position *p1* to the corresponding position *p1*' directly above/below *p1* on another floor (Preliminary Experiment 2 and Experiment 7), and (3) *paper-based drilling task*: the experimenter provides a floor layout (printed on a paper) and then asks participants to draw circles on the layout to indicate the horizontal locations of the targets located on another floor (Preliminary Experiment 2 and Experiment 7). The three vertical alignment tasks were used in different experiments of this thesis and play an important role in accessing users' formed between-floor alignment information. By definition, the drilling task is the most straightforward way to test vertical alignment information. Thus, for convenience, in this thesis the vertical alignment task refers to the drilling task unless explicitly stated otherwise.

We propose that multi-level survey knowledge is *sufficient* for supporting these cross-level spatial tasks. The term "sufficient" means that A (multi-level spatial knowledge), if satisfied, guarantees that B (cross-level spatial tasks) is obtained. In addition, we propose that multi-level survey knowledge is *necessary* for supporting these cross-level spatial tasks in a complex building. The term "necessary" means that A (multi-level spatial knowledge), must be satisfied in order for B (cross-level spatial tasks) to be obtained. Thus, we can use these cross-level spatial tasks to evaluate users' formed multi-level survey knowledge.

On the basis of the definition of multi-level survey knowledge, we argue that well-developed multi-level cognitive maps are not only useful for the obvious applications of affording efficient inter-level indoor route planning and wayfinding but that they could also be crucial for supporting many other scenarios. For example, in an emergency situation, firefighters need to determine the correct place to break through a ceiling to rescue people trapped in a building, or maintenance workers need to figure out the best route for drilling a hole to install conduit between floors. In each of these situations, accurately formed multi-level cognitive maps would be critically important for supporting spatial behaviors requiring integration of vertical spatial knowledge.

On the basis of the analysis of the logical relation between cross-level spatial tasks and multi-level survey knowledge, we conclude that, first, navigators need only a set of single-level survey knowledge representations and between-floor connectivity information to accurately and efficiently find targets located on different floors (between-floor wayfinding task). No between-floor alignment information or encoding of the z-axis is necessary in the between-floor wayfinding task. Second, in order to accurately point between positions located on different floors (point horizontally), navigators require only a set of single-level knowledge representations and between-floor alignment information. No between-floor connectivity information or encoding of the z-axis is needed in the between-floor pointing task. As discussed in Section 2.2.1, if navigators have learned the between-floor perspective shift, they need at least one pair of transition points (connectivity information) to calculate between-floor alignment information. Thus, under the prerequisite that users can accurately find positions located on different floors, the between-floor pointing task is the most important task for evaluating users' formed multi-level survey knowledge in this thesis. If a factor assists users to acquire more accurate multi-level survey knowledge, we say that this factor promotes users' development of multi-level cognitive maps.

2.5 Summary

This chapter first provided the core concept of this thesis, multi-level cognitive maps, referring to a globally coherent mental representation of multi-level built environments. We then described the framework of integrating cross-level spatial knowledge to develop a multi-level cognitive map. The main points of this framework include:

1) When a multi-level building has no between-floor visual access and no visual access between OI-spaces, navigators have to obtain rotational displacements by updating heading information during vertical transition (called path integration), in order to integrate cross-level spatial knowledge.

2) When a multi-level building has visual access to a global landmark, navigators can use configurations of landmarks to determine their location and direction (called piloting) and use the information to integrate cross-level spatial knowledge.

3) Secondary information sources such as AR visualizations and schematic maps that effectively convey the desired multi-level building information can facilitate the integration of cross-level spatial knowledge.

We then formally defined multi-level survey knowledge, which is the final product of the process of cross-level spatial knowledge integration. Finally, we designed three types of cross-level spatial tasks to test the development of multi-level cognitive maps. On the basis of these concepts and framework, we proposed three principal reasons for the challenge of developing multi-level cognitive maps including (1) complex between-floor structural and topological properties, (2) insufficient access to the requisite information such as global landmarks, and (3) the lack of cognitively motivated visualization techniques for indoor navigation. These assertions will be investigated throughout this dissertation in Experiments 1-7.

CHAPTER 3

GENERAL METHODS

In Chapter 3, we describe the general experimental methods that are in common across all experiments. Specific details are provided under each experiment.

3.1 Participants

All participants were recruited from the University of Maine student body. In total, 146 participants (73 females and 73 males) took part in all studies for this dissertation.

Table 3.1. Participant information.

Experiment	Females	Males	Mean (age)	SD (age)
Preliminary Experiment 1	10	10	20.9	2.0
Preliminary Experiment 2	6	6	25.9	6.7
Experiment 1	8	8	20.1	2.0
Experiment 2	8	8	21.6	1.8
Experiment 3	8	8	20.2	1.2
Experiment 4	8	8	20.1	2.0
Experiment 5	8	8	19.8	2.4
Experiment 6	8	8	21.4	4.4
Experiment 7	9	9	25.8	7.4

All participants self-reported as having normal or corrected to normal vision. All gave informed consent and received monetary compensation for their time.

3.2 Materials and Apparatus

The seven behavioral experiments in this thesis were carried out using Virtual Environments (VEs), as VEs and related technologies facilitate manipulation of building layout and information content as well as allow for tracking of navigators' movement behavior. In VEs, it is relatively easy to parametrically vary properties of the environment and to investigate how manipulating these environmental properties affect humans' mental representation (Loomis, Blascovich, & Beall, 1999), whereas in physical environments (PEs), controlled manipulations of floor layouts as well as landmarks would be very difficult, if not impossible, to similarly manipulate. Although there are several limitations of these VE-based systems, such as a smaller field of view, lower resolution, less realism than the PEs and often no auditory, tactile, proprioceptive and vestibular cues (Loomis, Blascovich, et al., 1999; Péruch & Gaunet, 1998), previous studies have shown that people can ultimately develop accurate spatial knowledge in large-scale VEs that is similar to knowledge acquisition gained from physical environments (O'Neill, 1992; Ruddle, Payne, & Jones, 1997; Stanton, Wilson, & Foreman, 1996; Tlauka & Wilson, 1996). Thus, VEs have been widely used by spatial cognition researchers to study spatial learning and wayfinding (Richardson et al., 1999; Tlauka et al., 2007; Vidal & Berthoz, 2005; Wiener & Mallot, 2003).

In this thesis, however, participants in all experiments had to navigate in a multi-level virtual building. It is necessary to establish a better understanding of the strengths and limitations of VEs on spatial learning and navigation in multi-level built environments. Richardson et al. (1999) required participants to learn the layout of a two-story building in one of three learning conditions, either from a map, from direct experience, or by traversing through a virtual rendition of the building. In the first testing phase, participants were required to make spatial judgments while imagining that they were within the environment (imagination condition)—imagining that they were directly facing each of the landmarks and then making direction estimates to other landmarks. In the second testing phase, the participants were led back into the real building for in-situ testing—the experimenter led participants to each landmark in the real building and asked for direction judgments to other landmarks. The results indicated that when pointing between landmarks on a single floor, VE learners and real-building learners showed similar performance with less than 5º difference in pointing error, suggesting that in a simple, single-floor environment navigators are able to "acquire as much knowledge from learning in a VE as from learning in a real environment" (Richardson et al., 1999). However, the VE learners had a greater between-floor pointing error than the real-building learners, and subsequent analysis revealed that this difference was mostly driven by the imagination condition, as the pointing error difference between VE and real-building performance was significantly larger during imagination responses (22º) than during in-situ responses (10º). Richardson et al. (1999) offered the interpretation that the real-building learners could look around and use what they could see of the building's layout to orient themselves, whereas the VE learners, especially in the imagination condition, did not receive self-orientation before testing (Richardson et al., 1999). To avoid this pitfall, in this thesis all cross-level spatial tasks (except the paper-based drilling task) were in-situ testing tasks, meaning that participants learned a virtual environment and were then tested in the virtual environment. In addition, all participants were encouraged to look around and use what they could see of the building's layout to become oriented before testing. These considerations should largely eliminate the between-floor pointing error caused by the VEs. Richardson et al. (1999) also pointed out that the VE learners had difficulty in correctly updating their heading while traversing the eight 90º turns in the stairwell. To address this issue, in this thesis participants used elevators to navigate between floors and thus only needed to make at most one turn during vertical transition (except in Experiment 7 and Preliminary Experiment 2 where participants navigated between floors using staircases).

In the study by Richardson et al. (1999), the virtual environment was displayed on a 15" monitor, and movement in the environment was initiated with the four arrow keys, enabling the user to look left and right or to move forward and back. In this thesis, we used a 43" monitor, and participants used a Logitech Extreme 3D Pro Joystick to execute both rotational and translational movement by rotating and pushing the joystick. We argue that physically rotating the stick about its z axis is more analogous to physical rotation and may facilitate the feeling of "presence" better than the keyboard interface used in Richardson et al. (1999)'s research. In addition, previous studies have found clear evidence that using larger visual displays (screen size) greatly increased the reported level of presence (Ijsselsteijn, de Ridder, Freeman, Avons, & Bouwhuis, 2001; Polys, Kim, & Bowman, 2007). In sum, we believe that, by avoiding the potential pitfalls that have been suggested as complicating vertical travel in VEs, combined with the continually improving quality and realism of VE renderings, similar cross-level spatial behavioral performance is obtainable for real and virtual navigation in multi-level built environments.

Additionally, we conducted two preliminary experiments to investigate the effects of the realism of the VE (sparsely rendered models vs. a photorealistic model), the immersion level of the VE (desktop VEs vs. head-mounted display VE), and the rotation method (physical rotation vs. imagined rotation) on users' development of multi-level cognitive maps. These factors are important for ensuring that the VE technology used in this thesis is the most valid measurement tool to address the research questions of interest.

Before introducing the two preliminary experiments, in the next section we will first introduce the general experimental procedure used in all experiments (specific details are presented under each experiment).

3.3 General Experimental Procedure

A within-subject design was adopted in all experiments. Generally, there were seven phases in each experiment.

Phase 1: *Practice*. Participants were familiarized with the apparatus and navigation behavior in the VE. All experimental tasks were explained and demonstrated before starting the experimental trials.

Phase 2: *Environmental learning*. In Experiments 1-6 and Preliminary Experiment 1, we used a method of environmental exposure and learning based on guidance from arrows on the floor. At the beginning of the experiment, participants were situated at one position either inside or outside of a computer-simulated multi-level building. A red arrow on the ground indicated north. Participants were asked to turn in place and to use what they could see of the environment (e.g., the building’s layout, the north arrow) to orient themselves. Participants were then guided by blue arrows on the ground to learn the entire building. When they passed by a target, an audio signal was played that indicated its name, such as “conference room.” In Experiment 7 and Preliminary Experiment 2, we used a method of free-exploration for environmental exposure and learning. With this technique, participants freely learned the building based on a user-defined open search method instead of being guided by arrows on the ground, as was done in the other studies. This modification was done as in these two studies we aimed to simulate “real” world navigation in which free exploration is perhaps the most common approach of environmental familiarization. However, in Experiments 1-6 and Preliminary Experiment 1, we had to ensure that users had built single-level survey knowledge in the learning phase, so the arrow-guided environmental learning method ensured similar coverage and visual information access across floors and between all participants. It is worth noting that in the Richardson et al. (1999) study, participants used a guided environmental learning method, whereas Ruddle et al. (1997) allowed participants to freely explore a single-level virtual building with an unlimited amount of time. The exploration was usually completed in 45 to 60 minutes.

Phase 3: *Pointing criterion task*. This task was designed to test whether participants had successfully learned all of the experimental targets in the building and had formed accurate single-level survey knowledge (2D cognitive maps). This task was critically important for ensuring that all participants had a common baseline level of learning and had formed accurate 2D cognitive maps. For this test, participants were first randomly situated at one target (e.g., a room) and a red arrow appeared to indicate north. The experimenter then asked them to get oriented by using any information they could avail themselves of by looking around the surrounding environment. When participants were ready, the experimenter asked them to turn to face a straight line to the elevator on the current floor as quickly as possible without compromising accuracy. To perform this task, participants rotated in the VEs by twisting the joystick and when they felt they were facing toward the elevator, pulled the trigger to log their response. A red crosshair on the screen indicated participants' facing direction. To meet the criterion, they needed to point to the elevator within a tolerance of 20° (15° in Preliminary Experiment 1). If they failed the first iteration, the Phase 2 learning and Phase 3 pointing criterion tests proceeded until they either successfully met criterion or made five incorrect attempts. About ten percent of the total participants failed this phase, partly due to experiencing motion sickness.

Phase 4: *Pointing task*. Participants were first randomly situated at a target and were told its name, e.g., "conference room". They were encouraged to orient themselves as they did in Phase 3. The experimenter then gave them a second target name and asked them to turn to face a straight line to that target (within-floor or between-floor). If the target room was on a different floor, they were instructed to ignore the vertical dimension and to point as if the target was on the same plane as their current floor. They pulled the joystick's trigger when they felt they were properly oriented so as to indicate a straight line to the requested target. Two dependent variables for the pointing task were analyzed: pointing latency and absolute pointing error.

Phase 5: *Wayfinding task*. Participants were first randomly positioned at a target and were told its name. They were encouraged to orient themselves as they did in Phase 3. The experimenter then gave them a second target name and required participants to navigate to the target (within-floor or between-floor) using the shortest possible route. Upon reaching the perceived location of the target, they turned to face it and pulled the joystick's trigger, at which point either the door opened (Experiments 1-6) or the target appeared (Experiment 7 and Preliminary Experiments 1-2) if they were correct. If incorrect, they were guided to the correct location before proceeding to the next trial. Two dependent variables were analyzed for this task: wayfinding accuracy (whether participants indicated the correct location and orientation of the target room) and wayfinding efficiency (shortest route length over traveled route length).

Phase 6: *Vertical alignment tests*. In Experiments 4-6, we used a *drilling task* to test learning of vertical alignment information. For this test, after participants had entered a room at the end of the wayfinding task (*Phase 5*), the experimenter asked them which room or object was directly above/below their current location. There were four options:

(1) a target room, (2) an empty room, (3) fire extinguishers or water fountains, and (4) nothing. The dependent variable for the drilling task was drilling accuracy (whether participants successfully indicated which room or object was immediately above/below their current location). In Experiment 7 and Preliminary Experiment 2, we tested knowledge of vertical alignment using a *vertical navigation task* and *a paper-based drilling task.* In the *vertical navigation task*, participants were randomly placed at one of the target locations and given its name. Their task was to navigate to the corresponding point in the environment that was directly above or below the target at which they were currently located. For example, if participants were located at a target chair on the ground floor, their task was to navigate to the corresponding point on the second floor that was directly above the chair. If the target was located at a place where there was no corresponding point on the other floor, participants were asked to navigate to the location that was closest to the corresponding vertically-aligned point. The dependent variable for the *vertical navigation task* was vertical navigation accuracy (whether participants successfully navigated to the corresponding vertically-aligned point). In the *paper-based drilling task*, participants were first given a paper which depicted the building's first floor layout. The experimenter then provided the names of the targets on the second floor. The task was to draw circles on the first floor layout to indicate the horizontal locations of the targets located on the second floor. Next, participants were given the second floor layout

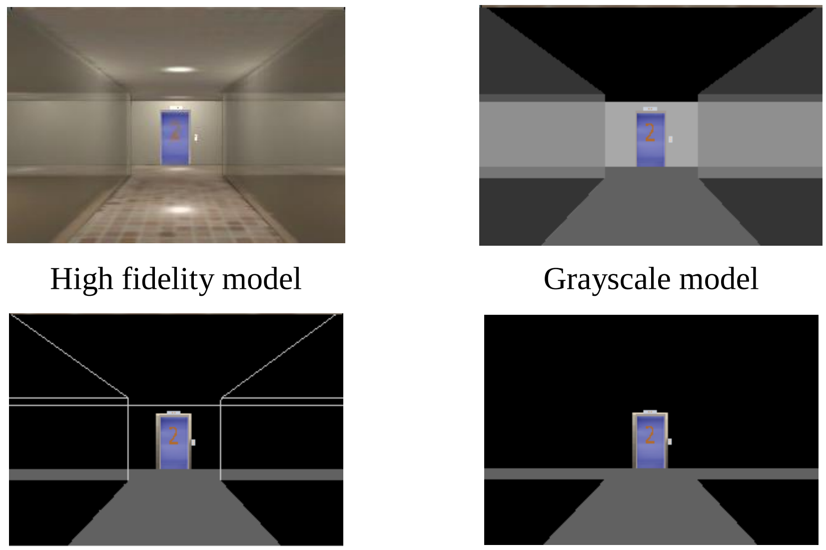
and asked to indicate the first floor targets using the same method. The floor order was balanced. In Experiment 7, each building had four targets, so there were four paper-based drilling trials for each building. In Preliminary Experiment 2, each building had three targets corresponding to three paper-based drilling trials for each building. Here the dependent variable for this task was the paper-based drilling accuracy (whether participants successfully indicated the cross-level target's location on the map).

3.4 Preliminary Experiment 1

The research question addressed by Preliminary Experiment 1 asks how the realism of the 3D visualization affects the development of multi-level cognitive maps. Empirical experiments addressing this issue support the view that users often have misplaced faith in realistic representations, termed "Naïve Realism" (Smallman & John, 2005). For example, people using spatial interfaces for naval applications prefer spatially realistic 3D icons of ships and planes on their displays vs. functional, symbolic icons. However, these realistic features were shown to actually decrease identification performance (Smallman & John, 2005). Similarly, users predicted they would need high-fidelity photorealistic 3D displays to find routes across outdoor terrain, whereas experimental results demonstrated that they actually performed the task better with lower fidelity displays (Smallman, Cook, & Cowen, 2007). Several studies have clearly shown that while photorealistic representations of maps appeal to users, they often have a negative impact on behavioral performance (Hegarty, Smallman, Stull, & Canham, 2009; Hegarty, 2008). As was illustrated in Klippel, Hirtle, & Davies (2010), people trying to use Google street view for wayfinding purposes converged on a similar experience—that simply providing photorealism is not enough for accurate spatial learning and wayfinding (Klippel et al., 2010). However, few studies have been conducted to evaluate the effect of environmental realism of 3D visualizations supporting real time indoor navigation. Although relatively impoverished renderings are assumed to be as effective in aiding people's navigation through indoor spaces as photorealistic models, this assumption has not been extensively studied, although initial evidence has provided some empirical verification. For example, Kalia, Legge, & Giudice (2008) found that richly rendered (photorealistic) indoor virtual models were not as efficient for spatial learning as a sparse model. However, Kalia et al. (2008)'s research did not investigate different levels of visual granularity of 3D visualizations, nor was it aimed at evaluating the efficacy of using a 3D visualization to facilitate development of multi-level cognitive maps, as is the goal here.

The primary goal of this experiment was to assess whether users' pointing and wayfinding performance after spatial learning with the 3D visualization differed as a function of the visual granularity of the interface, findings which will help specify the optimal information content to be used in future 3D displays for real-time indoor navigation systems. In this thesis, we aim to develop and evaluate perceptually-salient and user-friendly visualization interfaces to facilitate users' development of multi-level cognitive maps (investigated in Experiments 5-7). Thus, the findings of Preliminary Experiment 1 are important to the subsequent work of this thesis, as they ensure that the realism level of the 3D visualization we used in this thesis is effective to address the questions of interest.

In this study, we experimentally evaluated four simulation fidelity conditions which manipulate the level of visual granularity of the environment which is provided to the user by a simulated mobile device during learning of virtual buildings. Four levels of visualization granularity represent a natural progression of degraded surface detail for environmental rendering, while preserving building topology. Each model is depicted in Figure 3.1. The high fidelity model was rendered with photorealistic texture, natural light, and full color (The Mental Ray rendering plug-in was used to generate the model. The grayscale fidelity model used grey scale color to represent the building and there was no rendering of texture or photorealistic light. The wireframe model only rendered the lines at each edge. The sparse model was the simplest representation as it only contained the floor plan of each layout without walls and ceilings.



High fidelity model Grayscale model

Wireframe model Sparse model

Figure 3.1. Four visualization fidelity models.

3.4.1 Methods

Twenty participants (10 females and 10 males, mean age=20.9, SD=2.0) were recruited from the University of Maine student body.

We used an SX111 HMD (NVIS, Inc.), incorporating inertial tracking, a panoramic 111° field of view, and a high resolution 1260 x 1080 stereo display, which provides a highly immersive VR experience. Two Nintendo Wii remotes were used in the experiment. One was used by the experimenter to control the sequence of experimental phases, and the other was used by the participant to translate through the VE. Turning in the VE was done through physical body rotation. The Vizard 3D rendering suite, by

WorldViz Inc., was used as the VE platform supporting users' real-time navigation and recording their trajectory and test performance.

Our environments were comprised of five two level buildings which were richly rendered in the VE. 3DS Max was used as the 3D modeling and rendering tool. Each level of the building was based on a 3 x 3 matrix of hallways, as illustrated in Figure 3.2. Each hallway was subdivided into two corridor segments. We deleted two segments from the twelve possible corridor segments in the generic environment to create our experimental layouts. This procedure ensured that all the layouts were well matched in terms of number of nodes, segments, and intersections. The two floors were connected by two elevators, which also served as salient landmarks for orientation in each of the experimental buildings. From a top-down perspective, one elevator was always located at the top center and the other was located at the southeast corner, as shown in Figure 3.2.

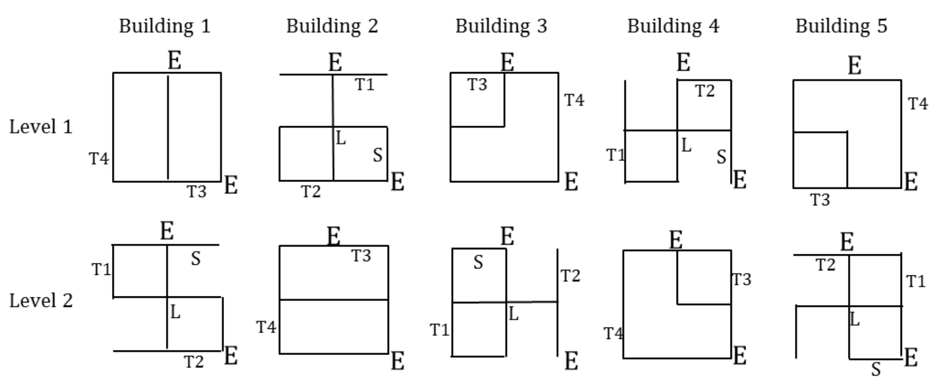


Figure 3.2. Experimental building layouts. (*E*) represents the elevators. (*L*) represents the starting position during the learning period, which was located at the only 4-way

intersection in the building. (*S*) represents the starting position for the wayfinding task. (*T*) represents the target.

The starting position for the wayfinding task was located near one of the two elevators to provide an orientation cue but was not visible from the starting learning point. There were two pictures on each floor which served as experimental targets. Pictures were based on eight high imagery words: bottle, chair, clock, dog, fish, kite, table and tie. All routes between pictures were matched across building for route length and number of turns.

A within-subject design was adopted, with the twenty subjects running in all five visualization conditions (high fidelity model, grayscale model, wireframe model, sparse model and a fifth unaided control condition). Participants first learned a multi-level building (*arrow-guided environmental learning*) and then took part in a pointing task and the wayfinding task. The pointing task in this study was different from the pointing task introduced in the general experimental procedure in that participants were asked to point from the learning start point (indicated by “L” in Figure 3.2), instead of pointing from a random target. Thus, the absolute pointing error in Preliminary Experiment 2 was smaller than that in the other experiments.

3.4.2 Results

A 5 (visual granularity: high fidelity model, grayscale model, wireframe model, sparse model and a fifth unaided control condition) × 2 (within-between floor: within vs. between) repeated-measures ANOVA was conducted for each of the four dependent measures of pointing latency, absolute pointing error, wayfinding accuracy and wayfinding efficiency. Significant main effects of within-between floor were observed for pointing latency, absolute pointing error, and wayfinding accuracy: pointing latency, $F(1, 39) = 10.79$, $p < .005$, $\eta^2 = 0.217$; absolute pointing error, $F(1, 39) = 6.495$, $p < .05$, $\eta^2 = 0.143$; wayfinding accuracy, $F(1, 39) = 9.457$, $p < .005$, $\eta^2 = 0.195$. The results showed that participants took longer times to point, exhibited greater pointing errors, and had lower wayfinding accuracy when pointing and wayfinding to targets located on different floors than when they were on the same floor.

Significant main effects of visual granularity were observed for both wayfinding accuracy and wayfinding efficiency: wayfinding accuracy, $F(4, 156) = 2.678$, $p < .05$, $\eta^2 = 0.064$; wayfinding efficiency, $F(4, 156) = 3.192$, $p < .05$, $\eta^2 = 0.076$. Subsequent Dunn–Sidak pairwise comparisons showed that navigation in the sparse model was more accurate than in the high fidelity model and the grayscale model, and navigation in the sparse model was more efficient than in the grayscale model (all *p*s < 0.01), suggesting that use of a sparsely rendered 3D visualization is more efficient than both of the high fidelity and grayscale models for assisting between-floor wayfinding performance. In addition, no reliable differences were found between pointing in the high fidelity model and in the sparsely rendered models, providing clear evidence that adding realism to the 3D visualization during learning is neither necessary nor advantageous for extraction of direction relations between targets.

3.4.3 Discussion

The primary goal of this experiment was to investigate whether a sparsely rendered model (e.g., the wireframe model and the sparse model) is as effective as a highly rendered model (e.g., the high fidelity model) for supporting the development of multi-level cognitive maps. The results showed that using the sparse model to assist learning led to the highest wayfinding performance and no reliable differences were found between pointing with the assistance of different visual granularity levels. These results provide compelling evidence that there is no reliable advantage of 3D visualizations rendered at a high level of visual granularity on learning and navigation of buildings and that in many cases, the best performance is obtained using a sparsely rendered spatial model. These findings are consistent with, and extend, previous research regarding the evaluation of the realism of 2D maps (Hegarty et al., 2009; Hegarty, 2008; Smallman et al., 2007; Smallman & John, 2005). One explanation is that participants need to extract picture and layout information from high fidelity 3D visualizations to encode the relative positions of these pictures as well as their positions in the building, whereas this information is more directly specified from the sparse model. This synthesis and extraction process may yield additional cognitive effort during learning which resulted in the increased navigation error and decreased efficiency for information-rich displays compared to the displays rendered with lower visual granularity. As illustrated by Smallman & John (2005), good display design is more than slavishly adhering to realism. Our research extends the theory of naïve geography to use of 3D visualization in real time indoor navigation and provides new evidence for the basic principle of these displays that graphics should not provide more information than is needed by the user (Smallman & John, 2005). The findings of Preliminary Experiment 1 set the stage of the model rendering requirements needed in Experiments 5-7 and ensure that a photorealistic model is not necessary and might be less effective for assisting the development of multi-level cognitive maps.

3.5 Preliminary Experiment 2

The first research question addressed by Preliminary Experiment 2 asks how the immersion level of the virtual environment affects the development of multi-level cognitive maps. Immersion "describes the extent to which computer displays are capable of delivering an inclusive, extensive, surrounding, and vivid illusion of reality to the senses of human participants" (Slater & Wilbur, 1997). This research question is addressed by comparing highly immersive VEs where the simulated information is presented through a head-mounted display (HMD) and changes with head movement vs. low immersion desktop VEs, where the simulated information is presented on a computer monitor. Compared to desktop VEs, immersive VEs generally have stereoscopic vision with a wider field of view and tracking of more degrees of freedom. Navigators can immerse themselves in the VEs to obtain realistic views and a sense of presence which is not possible using desktop VEs (Loomis, Blascovich, et al., 1999). Therefore, there is a common assumption that they are more effective than desktop VEs for certain spatial behaviors (e.g., pointing and wayfinding). However, immersive VE equipment, e.g., the HMD and head-tracking sensors, are far more expensive and complex to set up than desktop VE systems (Wright & Madey, 2008) and are therefore rarely used outside of the research lab setting. In addition, previous literature has shown that the HMD-based systems often elicited higher levels of simulator sickness and evoked larger negative emotions compared to desktop VEs (Kim, Rosenthal, Zielinski, & Brady, 2012). By contrast, desktop VEs have become less expensive and the graphics cards have become more powerful, so viewing and interacting with high quality desktop VE environments is more easily managed than in the past (Lau et al., 2003). Hence, if desktop VEs are shown to lead to similar behavioral performance as is obtained from use of an HMD in some circumstances, use of this simpler technology may open the door for their implementation in a much broader application domain by many more people, instead of being limited to a few research labs as is the case with immersive VE systems. Thus, it is necessary to assess whether there is sufficient benefit to justify using the more expensive immersive VE systems in spatial cognition research. Some evidence from previous literature supports the use of desktop VE systems. For instance, Ruddle, Payne, & Jones (1999) asked participants to walk through a virtual building using an HMD and the same environment using a desktop VE. Results of the study showed that there was no significant difference in the absolute percentage error of participants' straight-line distance estimates; also, there were no reliable differences in the direction-estimates between the two types of displays (Ruddle et al., 1999). However, the simulated environments used in Ruddle et al. (1999)'s study were two single-level virtual buildings. It is unknown from previous literature how the immersion level of the virtual environment affects the development of multi-level cognitive maps. This issue is addressed in the current experiment.

The second research question addressed in this experiment asks whether the rotation method used in the desktop VE (imagined rotation vs. physical rotation) affects users' development of multi-level cognitive maps. Navigation is a common interactive task performed with spatial cognition studies in VEs. However, the tracked space in the physical world is usually much smaller than the VE being navigated. In immersive VEs, previous techniques either replicate the motions of walking (e.g., treadmills, walking in place) or employ a joystick or keyboard to effect translation, while direction of movement is usually specified by either head orientation or a handheld pointer. In desktop VE systems, the joystick or keyboard is often used for both translation and orientation. Riecke et al. (2010) conducted a study in which participants were asked to search through a computer generated environment for targets in a joystick condition, free walking condition (physical translational and rotational movement) and rotation condition (only rotational movement). In the joystick condition, both horizontal translations and yaw/pitch rotations in the VE were controlled by the joystick. Participants in the walking condition navigated through the virtual scene by physically walking (translating and turning). Participants in the rotation condition used a joystick to translate through the virtual scene, but rotations were still controlled by corresponding physical motions (turning in place). The results showed that physical rotations alone without actual walking are sufficient for supporting users in finding targets in a single-level VE using an HMD. In addition, Giudice & Tietz (2008) conducted a study to investigate the effect of physical body rotation using virtual verbal displays (without HMDs) for environmental learning and wayfinding based on spatial language vs. visual access. Their results also showed that employing physical rotation during learning significantly improved spatial knowledge acquisition and cognitive map development (Giudice & Tietz, 2008). Although previous literature has shown that translation with HMDs or virtual verbal displays based on a joystick as the means of movement through a single-level virtual space is effective, the effectiveness of physical rotation in a multi-level desktop VE on users' development of multi-level cognitive maps is still unknown. If the results show that imagined rotation in a desktop VE is effective in supporting multi-level cognitive map development, the setup of desktop VEs will be further simplified, as experimenters do not need to utilize an inertial tracking system to update users' heading in the physical rotation conditions.

In the current experiment, three within-subject conditions (physical rotation HMD, physical rotation desktop and imagined rotation desktop) were used. In the physical rotation HMD condition, participants wear an HMD to apprehend the space as they navigate the VEs. A joystick is used to perform forward translation and rotations are made by spinning in place on a chair. An inertial tracker is used to update users' change in heading with rotation. In both of the two desktop conditions, participants used a laptop to see the virtual building. Similar to the physical rotation HMD condition, navigation in the physical rotation desktop VE was done using a joystick to perform translational movement and physical turning via the rotating chair was used to execute rotations.

However, in the imagined rotation desktop condition, rotations were executed by twisting the stick about its z axis, which is the only difference between the two desktop conditions.

3.5.1 Methods

Twelve participants (6 females and 6 males, *M* =25.9, *SD*=6.7) were recruited from the University of Maine student body.

In the immersive HMD VE condition, we used a zSight integrated SXGA HMD (Sensics, Inc.), incorporating inertial tracking, 70° field of view, and a high resolution full-color SXGA 1280×1024 pixels per eye. A Lenovo W510 Thinkpad 15.6-inch workstation notebook with an Intel Core i7 processor and NVIDIA Quadro FX 880M graphics was used in the two desktop VE conditions. A Logitech Extreme 3D Pro Joystick was used in all three conditions to perform forward-back translations. In the imagined rotation desktop VE condition, participants use the joystick to make both translational and rotational movements. In all conditions, participants sat on a rotatable chair with an attached platform to hold the laptop used in the two desktop conditions.

Our environments were comprised of three two-level buildings which were designed using Revit Architecture 2013 (AutoDesk, Inc.), as shown in Figure 3.3. The

Unity 4.0 VR engine (Unity Technologies) was used as the VE platform supporting users' real-time navigation and recording their trajectory and test performance.

Figure 3.3. Multi-level virtual buildings.

All the virtual buildings used in the experiment were matched for layout complexity and topology, as shown in Figure 3.4. The start learning point is located at the southeast corner of the first floor. There was a red arrow indicating the start point and the north direction. The two floors were connected by one staircase, located at the north of the building. The staircase and the start point arrow could serve as landmarks for orientation in each of the experimental buildings. There were two pictures on the second floor and one picture on the first floor which served as experimental targets. Pictures were based on three high imagery words: chair, fish and kite. All targets were initially hidden from view but when participants passed the target, an audio signal was triggered that gave its name. The target also visually appeared for ten seconds and then faded out.

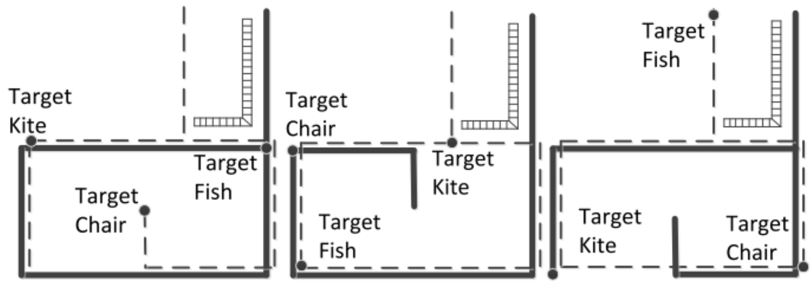


Figure 3.4. Experimental layouts with target locations denoted. The solid line represents the first floor layout, while the dashed line represents the second floor layout.

A within-subject design was adopted, with the twelve participants running in all three conditions (physical rotation HMD, physical rotation desktop and imagined rotation desktop). Participants first freely learned a multi-level building and then took part in four spatial tasks: pointing, wayfinding, vertical navigation task, and paper-based drilling tasks (as described in the general experimental procedure).

3.5.2 Results

A 3 (VE conditions: physical rotation HMD, physical rotation desktop and imagined rotation HMD) × 2 (within-between floor: within vs. between) repeated-measures ANOVA was conducted for each of the four dependent measures of pointing latency, absolute pointing error, wayfinding accuracy and wayfinding efficiency. There was no significant effect of VE conditions on any of these measures (all ps > .05), suggesting that more immersive VE systems (e.g., HMD) do not necessarily lead to better performance in the two cross-level spatial tasks. As with the previous experiment, significant main effects of within-between floor were observed for pointing latency, absolute pointing error, and wayfinding accuracy, with within-floor pointing trials being faster and more accurate than between-floor pointing trials: pointing latency, $F(2, 22) = 8.209$, $p < .05$; absolute pointing error, $F(2, 22) = 36.496$, $p < .0001$. Within-floor wayfinding was also more accurate than between-floor wayfinding: wayfinding accuracy, $F(2, 22) = 5.077$, $p < .05$. These findings are consistent with the findings of Preliminary Experiment 1. The findings of the two preliminary experiments motivated the research of this thesis—why is integrating between-level building information so challenging for human spatial cognition? This fundamental and core research question is investigated throughout the remainder of the thesis.

A one-way repeated measures ANOVA was run for each of the two dependent measures of vertical navigation accuracy and paper-based drilling accuracy with the three VE conditions as a within-subject factor. The main effect of VE conditions on paper-based drilling accuracy was significant, $F(2, 70) = 3.237$, $p < .05$. Subsequent Dunn–Sidak pairwise comparisons showed that average paper-based drilling accuracy in the desktop VE was significantly higher than that in the HMD VE, $t(35) = -2.256$, $p < .05$, and that the average drilling accuracy in the physical rotation desktop condition was significantly higher than in the imagined rotation desktop condition, $t(35) = -2.092$, $p < .05$.

No effect of VE conditions on vertical navigation accuracy was observed, $F(2, 70) = .836$, $p > .05$.

3.5.3 Discussion

The primary goal of this experiment was to investigate whether the desktop VE is as effective as more immersive HMD-based VEs for supporting the development of multi-level cognitive maps. The results showed that there were no reliable differences between the two VE immersion levels (HMD vs. desktop) in the pointing, navigation, vertical navigation and paper-based drilling tasks. There may still be advantages to immersive VEs, but the results of this experiment indicate that the benefit gained from more immersive VEs (e.g., HMD) may not be as pervasive as is suggested in the literature, at least for performing these cross-level spatial tasks. The most likely reason for the lack of an effect for HMD-based immersive VE is that the HMD VE cannot provide a vivid illusion of the vertical travel during this task. Thus, a higher immersion VE did not provide a benefit over the desktop VE. In addition, participants in the immersive condition needed to wear a heavy HMD for about 15 minutes (including the learning and testing phases) which sometimes has been shown to cause discomfort for participants (Kim et al., 2012). In corroboration, several participants in this experiment self-reported dizziness caused by wearing the HMD. As a result, the advantages of higher immersion are unfortunately offset by participants' dizziness and potential distraction from the VE equipment.

With regard to the rotation method (physical vs. imagined) in desktop VEs, the results showed that there was no significant difference between physical rotation and imagined rotation in the pointing, navigation and vertical navigation tasks. The only task where participants showed a reliable benefit from physical rotation was the paper-based drilling task. However, the pointing error for physical rotation was much higher than the other two conditions, so making any definitive claims for an advantage of inclusion of physical rotation is not possible from these data. In this experiment, users sat on a rotatable chair and thereby the rotation was less intuitive and natural than with the rotation in place while standing as was used in the virtual study with verbal displays where it was shown to be an advantage (Giudice & Tietz, 2008) and the virtual navigation study (Riecke et al., 2010). Therefore, the advantages of physical rotation, e.g., proprioceptive and vestibular feedback, may have been less in the current design. In addition, during vertical transition in the stairwell, participants had to turn smoothly when they went forward and upward. Several participants in the physical rotation conditions had difficulty in navigating smoothly along the staircases because they could not make the physical rotation and imagined joystick translation simultaneously. As a result, these participants usually made a turn too early or too late and thereby got stuck by the stair railing. However, in the imagined rotation desktop condition, users only needed to use one device to execute both rotational and translational movement behaviors by rotating and pushing the joystick. The vertical travel, providing important translation and rotation information, is unfortunately the hardest part for the virtual navigation in the physical rotation desktop VE, so the performance on the test measures in the imagined rotation desktop condition (with joystick) was no worse and perhaps even better than the physical rotation desktop VE condition.

In sum, some situations may benefit from higher immersion or physical movement but at least for the cross-level spatial tasks used in the current multi-level environments, the current data suggests that employing desktop VEs with a joystick to perform both translational and rotational movement is sufficient for studying human navigation in a virtual building and multi-level cognitive map development. Therefore, in the seven primary experiments described in this thesis, participants used desktop VEs (with a Logitech Extreme 3D Pro Joystick) to translate and rotate in the virtual environments.

3.6 Summary

This chapter first described the general methods pertaining to the majority of the experiments elaborated in this thesis. We then described two preliminary experiments that were conducted to examine basic VE characteristics that set the stage for subsequent experiments described in later chapters. This included evaluating the effects of the realism of the VE, the immersion level of the VE, and the rotation method used in the VE on users' development of multi-level cognitive maps. In Preliminary Experiment 1, we found that 3D visualizations rendered at a high level of visual granularity are not necessary for effective learning and navigation of multi-level buildings. Preliminary Experiment 2 indicated that using a desktop VE and employing a joystick for both translation and rotation is sufficient for studying human navigation in a virtual multi-level building. These studies refining the experimental tool (VEs) for supporting navigation in multi-level built environments are important to the subsequent work of this thesis, as they help determine the technical requirements for all seven behavioral experiments and ensure that the tool we chose (desktop virtual reality) is effective to address the questions of interest. Importantly, the findings of the two preliminary experiments (the between-floor effect) were the impetus to explore why integrating between-level building information is so challenging for human spatial cognition. This fundamental research question is investigated throughout the following experiments.

CHAPTER 4

EVALUATION OF ENVIRONMENTAL FACTORS ON THE DEVELOPMENT OF MULTI-LEVEL COGNITIVE MAPS

4.1 Introduction

In Chapter 4, we describe four experiments (Experiments 1-4) using VEs to evaluate the impact of environmental factors including (1) multi-level structural and topological properties of built environments, and (2) global indoor and outdoor landmarks in the environment, on the development of multi-level cognitive maps. In Section 2.2.1, we argued that between-floor structural and topological properties of buildings impair users' development of multi-level cognitive maps (*Hypothesis 1*). To evaluate this hypothesis, Experiments 1-3 investigated five between-floor structural and topological properties of buildings: (1) the z-axis offset, (2) the 90° between-floor heading shift, (3) the between-floor overlap, (4) the between-floor misalignment, and (5) the portal-floor heading shift. Each of these properties will be described below and are illustrated in Figure 4.1 and Figure 4.2.

(1) *The z-axis offset*. A multi-level building contains multiple floors and each floor has a z-axis value (e.g., floor height), meaning that different floors have a *z-axis offset* (i.e., vertical distance between floors), as shown in Figure 4.1(b). To study the effect of

the z-axis offset, we designed a single-floor building with two regions, pictured in Figure 4.1(a), serving as the control condition. As discussed in Section 2.2, a *region* represents a perceived and encoded representation in spatial memory in which locations are grouped within a common spatial reference frame (Wiener & Mallot, 2003; Wiener et al., 2005). Note that, in this thesis, the two regions of the single-floor building (the control condition) were connected by an “elevator”, which supported users’ horizontal transition between the two regions on the same plane instead of going up/down between floors. Importantly, the two regions of the single-level building (no z-axis offset, control condition) and the two floors of the multi-level building (including a z-axis offset) were matched for layout complexity. Therefore, by comparing users’ performance on cross-level spatial behaviors (see description in Section 2.4) between the two types of buildings, we can examine the effect of the z-axis offset on users’ development of multi-level cognitive maps.

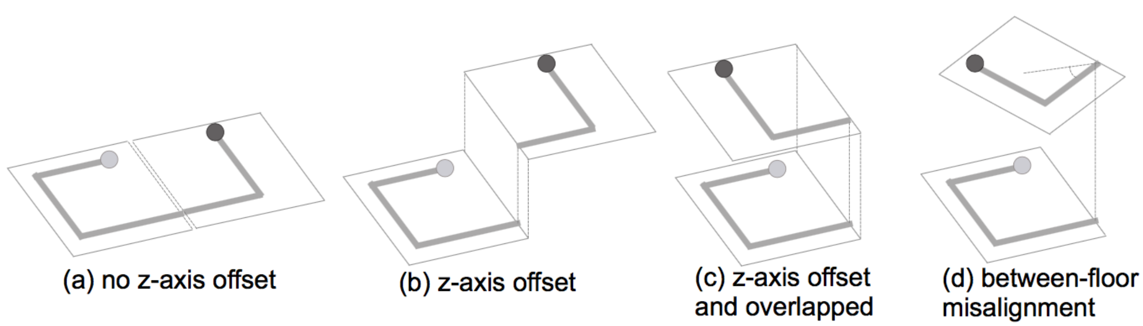


Figure 4.1. The z-axis offset, between-floor overlap, and misalignment. The bold solid lines represent the hallways.

(2) *The between-floor overlap*. Two floors of a building can be displaced so as to be non-overlapped between the floors, as in Figure 4.1(b), or be overlapped between the floors (Figure 4.1(c)). This can be thought of as a continuum from no overlap (0%) to completely overlapped (100%), if two floors/regions are matched at their boundaries, as shown in Figure 4.1 (see discussion in Egenhofer, 1993). If two floors of a building are overlapped, there must be a set of positions within the two floors co-located at the same x-y coordinates. For simplicity in this thesis, however, between-floor partially- and fully-overlapped floors are both termed as encompassing the *between-floor overlap* factor. The examination of different types of between-floor overlap on multi-level cognitive map development is beyond the scope of this thesis but an important research topic for future studies.

(3) *The between-floor misalignment*. If two floors of a building have an angular offset (i.e., perspective shift γ, see Section 2.2.1), they can be said to have a *misalignment* between the floors, as shown in Figure 4.1(d), known as the *between-floor misalignment* factor.

(4) *The between-floor heading shift*. As described in Section 2.2.1, navigators use vertical connectors (e.g., elevators, staircases, escalators) to navigate between (up/down) floors. In this process they pass through a between-floor portal such as an elevator/stairway door to enter or exit a floor. The *between-floor heading shift* refers to

the heading shift (if present) imposed when entering and exiting between-floor portals (see Section 2.2.1). For example, as shown in Figure 4.2(b), the elevator has separate entrance and exit doors that are offset by 90°, so navigators will experience a 90° between-floor heading shift after vertical travel.

(5) *The portal-floor heading shift*. If the orientation of the portal has an angular offset with respect to the floor's reference direction, navigators will experience a heading shift both before entering and after exiting the portal. This heading shift is termed the *portal-floor heading shift* (see Section 2.2.1), referring to the angular offset between the reference direction of a floor and the orientation of a portal. For example, as shown in Figure 4.2(c), navigators will experience not only a 90° between-floor heading shift after vertical travel but also a 45° portal-floor heading shift before entering and after exiting the elevator.

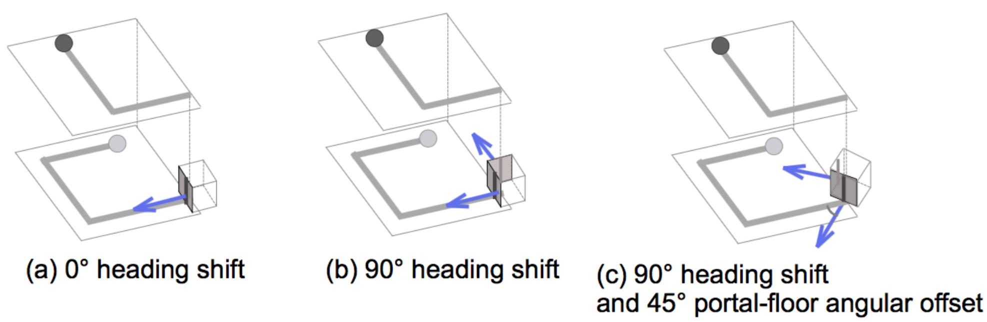


Figure 4.2. Between-floor heading shift and portal-floor angular offset. The blue arrows indicate the orientation of the elevator's entrance and exit doors.

We studied these five between-floor structural and topological properties in Experiments 1-3. In addition, Experiment 4 examined the effects of global landmarks (both indoor and outdoor) on users' development of multi-level cognitive maps in order to evaluate the two hypotheses proposed in Section 2.4: (1) visual access to a global landmark from within a building promotes users' development of multi-level cognitive maps (*Hypothesis 2*) and (2) if a global landmark is visible from both indoor and outdoor spaces, it will facilitate both cross-level spatial knowledge integration and the integration of OI-spaces (*Hypothesis 3*).

4.2 Experiment 1 (z-axis offset and 90° heading shift)

4.2.1 Introduction

The first research question addressed by Experiment 1 asks whether the z-axis offset can solely be attributed as impairing the development of multi-level cognitive maps. In the two preliminary experiments, we found that participants took longer to point, exhibited greater pointing errors, and had lower wayfinding accuracy when pointing and wayfinding to targets located on different floors than when they were on the same floor. However, in these two studies, the between-floor routes were more complex than the within-floor routes with more turns and longer overall route length. In order to accurately point between locations without direct visual access, navigators have to integrate local spatial knowledge based on turns taken, distances, and angles (Golledge, 1999). However, errors can accumulate during this process—the longer the travelled route, the larger the path integration error (Etienne & Jeffery, 2004; Klatzky, Loomis, & Golledge, 1997; Loomis et al., 1993a; Wan, Wang, & Crowell, 2013). In addition, the multi-level buildings used in the two preliminary experiments consisted of two overlapped floors. Thus, it is unclear whether the effects of the between-floor pointing and navigation performance found in these two studies were caused by the longer travelled routes, the z-axis offset, or the between-floor overlap (or a combination thereof that cannot be disentangled). Due to the nature of the experimental design, the majority of previous studies on indoor wayfinding in the real world also suffer from difficulty determining which factor (or factors) cause the between-floor effect of pointing and wayfinding. It is not surprising that previous literature investigating the effects of between-floor pointing performance is somewhat contradictory. For instance, Montello and Pick (1993) conducted a study with results suggesting that the z-axis offset factor did not contribute to the difficulty of between-floor pointing. In their study, participants first learned two distinct routes in and around a university building that never crossed. One route used in their study contained two sections with a z-axis offset, meaning one section was inside the building whereas the other was outside. The outside section of the route was one floor above the inside section. After walking the route and learning object locations (not visible from each other), participants were required to indicate straight-line directions to landmarks in the two sections (located on two levels) to test the formed cognitive map. The results showed that between-level pointing accuracy was comparable, suggesting that the z-axis offset did not affect users' pointing performance. However, the within- and between-floor routes used in Montello and Pick (1993)'s study had a different path structure in terms of the number of turns and overall route length. Thus, the relative pointing performance was difficult to interpret, and it is still unknown whether the z-axis offset alone impairs the development of multi-level cognitive maps.

To address this issue, in this thesis we designed a set of "ideal" environments using VEs to disentangle factors discussed above that may lead to the between-floor effect. We aimed to determine exactly which topological or structural factor of multi-level built environments leads to the difficulty in between-floor pointing and wayfinding known from the literature and from our preliminary experiments described in Chapter 3. For instance, in Experiment 1, we designed four non-overlapped buildings using VEs, depicted in Figure 4.3. Buildings 3 and 4 had two non-overlapped floors (but with a z-axis offset), whereas buildings 1 and 2 had only one floor with two non-overlapped regions (no z-axis offset). Importantly, the overall route length and the number of turns for within- and between-floor routes were matched for each building. In this case, we successfully excluded the two structural factors: the longer-travelled routes and the between-floor overlap. Therefore, if an effect of between-floor pointing or wayfinding performance was observed in Experiment 1, we can conclude that the z-axis offset solely impairs the development of multi-level cognitive maps (see discussion of cross-level spatial tasks in Section 2.4).

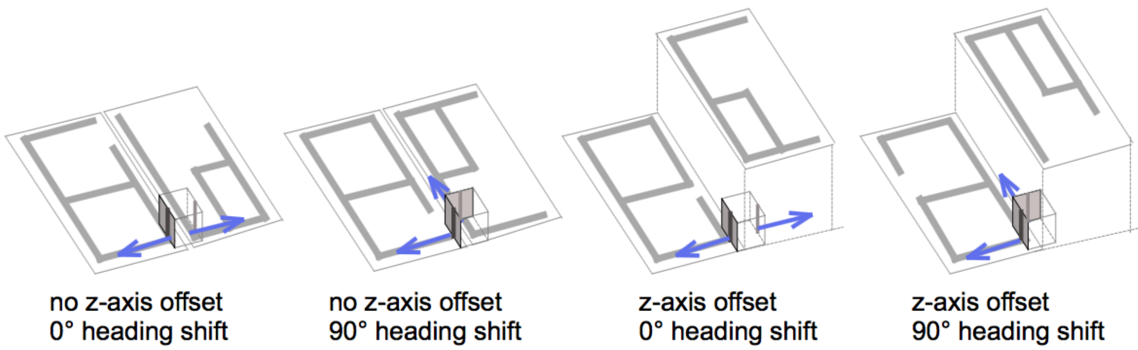


Figure 4.3. Floor layouts of Experiment 1.

The second research question addressed by Experiment 1 asks whether the 90° heading shift during between-floor or between-region transitions impairs the development of multi-level cognitive maps. In Section 2.2.1, we proposed that the between-floor heading shift during vertical transition might cause the difficulty of integrating multi-level spatial knowledge. This assertion is investigated in the current study. Previous literature has found that confusing staircases are one of the main reasons for becoming lost inside buildings (Hölscher et al., 2006). However, two factors are involved in the vertical transition via a staircase: the between-floor heading shift (as discussed in Section 2.2.1) and additional movements and turns in the stairwell. Thus, it is unclear whether the difficulty of using staircases for vertical transition found in the research by Hölscher et al. (2006) was caused by the between-floor heading shift or by the additional movements and turns in the stairwell. To address this issue in Experiment 1, navigators only used elevators for vertical transitions, as this eliminates any potential confound from the additional rotations imposed by using stairs. In addition, we designed two types of elevators, one with and one without the between-floor heading shift, as shown in Figure 4.3. Buildings 1 and 3 had an elevator with a 0° heading shift, whereas buildings 2 and 4 had an elevator with a 90° heading shift. Given that there are no additional movements and turns when using the elevator for vertical transitions, if an effect of between-floor pointing or wayfinding performance is observed, we can conclude that the 90° between-floor heading shift impairs the development of multi-level cognitive maps.

In research by Street (2012) on indoor navigation, two groups of participants learned a multi-level campus building. One group used an elevator to navigate between floors, and the other group used a staircase for vertical transition. After learning the building, both groups were asked to point to within-floor and between-floor targets (with the same route length). The results showed that the overall pointing error for the group using the elevator was significantly less than of the group using the staircase. However, for the elevator group, navigators still had larger between-floor pointing error than in the within-floor pointing trials, suggesting that any additional movements and turns imposed by use of the stairs for between-floor transitions cannot be the sole source of the difficulty observed in between-floor pointing performance. Street (2012) did not propose what factor might be causing this between-floor effect. However, we postulate that it was likely caused by the combination of the z-axis offset and the between-floor overlap factors as described in the introduction. This assertion is further investigated in Experiment 2.

4.2.2 Methods

Sixteen participants (eight females and eight males, mean age = 20.1, *SD* = 2.0) were recruited from the University of Maine student body.

The experimental environments were displayed on a Samsung 43" Class Plasma HDTV monitor running at 60 Hz at a resolution of 1024 × 768, as shown in Figure 4.4. We ran the desktop VEs using a Lenovo W510 Thinkpad 15.6-inch workstation notebook (Intel Core i7 processor and NVIDIA Quadro FX 880M graphics). We used the Unity 4.0 VR engine (Unity Technologies) as the VE platform supporting users' real-time navigation and recording their trajectory and test performance. Participants used a Logitech Extreme 3D Pro Joystick to make both translational and rotational movements (see discussion in Preliminary Experiment 2).

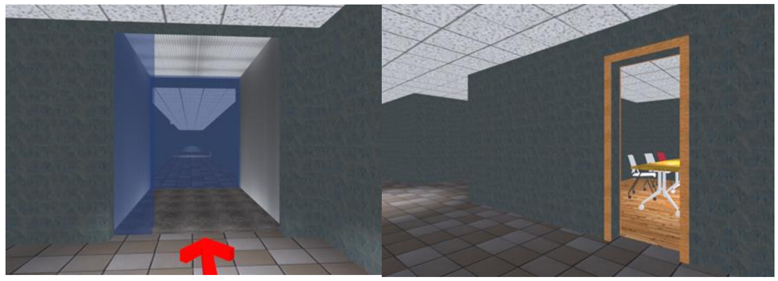


Figure 4.4. Virtual Environments of Experiments 1-3.

Our environments comprised four buildings designed using Revit Architecture 2013 (AutoDesk, Inc.). The four buildings were matched for layout complexity but had distinctive between-floor topological and structural properties (i.e., the z-axis offset and between-floor heading shift), as pictured in Figure 4.3. Each building contained an elevator and four rooms: a bathroom, a classroom, a conference room, and an office, serving as targets. The four rooms had the same size (5m×5m) but distinctive interior objects and floor textures. The locations of the four rooms were balanced among the four buildings. As described above, the overall route length and the number of turns for within- and between-floor routes were matched for each building.

All participants of Experiments 1-3 followed the same procedure across experiments: they first learned a multi-level building and then took part in two cross-level spatial tasks: between-floor pointing and wayfinding (see Sections 2.4 and 3.3). The vertical alignment task was not used in Experiment 1, as all buildings were non-overlapped.

4.2.3 Results

A 2 (z-axis: no offset vs. offset) × 2 (between-floor heading shift: 0° vs. 90°) repeated-measures ANOVA was conducted for each of the four dependent measures (absolute pointing error, pointing latency, wayfinding accuracy, and wayfinding efficiency). There was no significant main effect of the z-axis offset for any measure: absolute pointing error, $F(1, 63) = 0.081$, $p > .05$, $\eta^2 = .001$; pointing latency, $F(1, 63) = .336$, $p > .05$, $\eta^2 = .005$; wayfinding accuracy, $F(1, 31) = 0.177$, $p > .05$, $\eta^2 = .006$; or wayfinding efficiency, $F(1, 31) = 0.140$, $p > .05$, $\eta^2 = .004$.

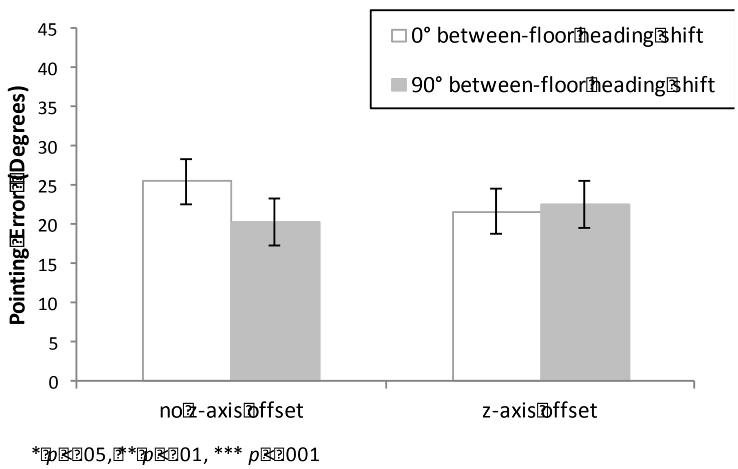
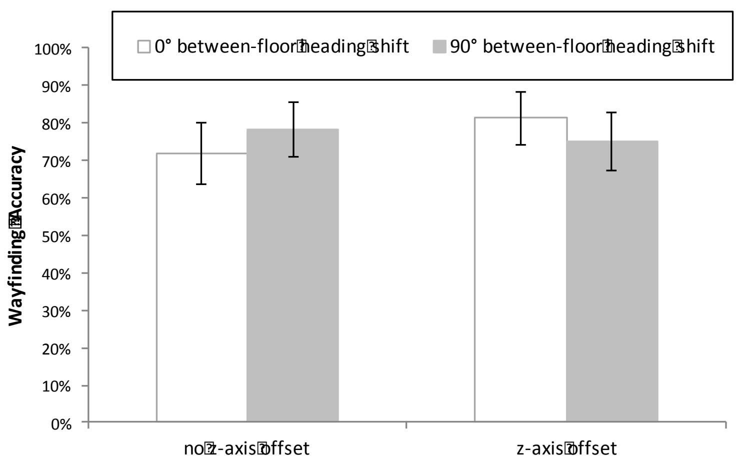


Figure 4.5. Mean absolute pointing error for Experiment 1.

Note that in this study the between-floor and within-floor routes had exactly the same number of turns and equivalent route length. If the z-axis factor led to the difficulty of cross-level spatial knowledge integration, participants must exhibit greater errors when pointing and wayfinding to targets located on different floors than when they were on the same floor. However, no effect of the z-axis offset was observed for any measure. Based on the observed effect sizes, we used G*power to determine how many subjects would have been necessary for statistical significance at power level of .95. We found that statistical significance was unlikely unless sample size was dramatically increased (n > 100,000). Thus, we are confident in the null results. In other words, the z-axis offset alone cannot be attributable to impairing users' development of multi-level cognitive maps.



*p < .05, **p < .01, *** p < .001

Figure 4.6. Mean absolute wayfinding accuracy for Experiment 1.

There was no significant main effect of the 90° between-floor heading shift for any measure: absolute pointing error, $F(1, 63) = .574$, $p > .05$, $\eta^2 = .009$, pointing latency, $F(1, 63) = .164$, $p > .05$, $\eta^2 = .003$, wayfinding accuracy, $F(1, 31) = 0.0001$, $p > .05$, $\eta^2 = .0001$, wayfinding efficiency, $F(1, 31) = 0.001$, $p > .05$, $\eta^2 = .0001$. The results suggest that the 90° between-floor heading shift also did not impair users' development of multi-level cognitive maps, at least not in any of the Experiment 1 buildings. Our lack of an effect may be due to our building design; that is, all of the Experiment 1 buildings were non-overlapped but shared the same reference direction, so participants might have used this as a cue. This issue will be further studied in Experiment 3, in which we investigated whether a confusing heading shift (the combined factor of between-floor heading shift and portal-floor heading shift) would impair users' development of multi-level cognitive maps in overlapped and aligned buildings.

Table 4.1. Mean absolute pointing error, pointing latency, wayfinding accuracy, and wayfinding efficiency in Experiment 1.

Dependent Variables (Measures)	Independent Variable II	Independent Variable I	
		no z-axis offset	z-axis offset
Absolute pointing error	0° between-floor heading shift	25.39 (2.88)	21.58 (2.89)
	90° between-floor heading shift	20.29 (3.05)	22.48 (2.96)
Pointing latency	0° between-floor heading shift	12.48 (0.83)	11.85 (0.90)
	90° between-floor heading shift	12.52 (0.83)	12.32 (1.13)
Wayfinding accuracy	0° between-floor heading shift	71.9% (8.1%)	81.3% (7.0%)
	90° between-floor heading shift	78.1% (7.4%)	75.0% (7.8%)
Wayfinding efficiency	0° between-floor heading shift	70.6% (8.0%)	79.9% (6.9%)
	90° between-floor heading shift	77.4% (7.4%)	73.5% (7.7%)

4.2.4 Discussion

The primary goal of Experiment 1 was to investigate whether either a z-axis offset or a 90° between-floor heading shift impairs the development of multi-level cognitive maps. Interestingly, no effects of the two factors were observed for any measure.

In Experiment 1, navigators learned each floor/region separately whether or not there was a z-axis offset, meaning that the z-axis offset factor only affects the encoding of the vertical dimension not the other three components of multi-level cognitive maps (multiple single-level 2D cognitive maps, between-floor connectivity, and alignment information). According to the logical relation between multi-level cognitive maps and cross-level spatial tasks, the encoding of the z-axis offset is not necessary for the between-floor pointing and wayfinding tasks (see Section 2.4). Thus, the lack of an effect of the z-axis offset on the between-floor pointing and wayfinding performance is consistent with the framework of multi-level cognitive map development introduced in Sections 2.1 through 2.4. An implication of this finding is that some multi-level buildings such as stadiums and terrace-like architectures, although comprising multi-level structures, are not necessarily more challenging to learn and to navigate than their single-level counterparts. Although the effect of the z-axis offset was not observed in Experiment 1, it does not rule out an effect from the z-axis. It simply indicates that the z-axis offset alone is not inherently the problem. In the two preliminary studies (Giudice

& Li, 2012; Li & Giudice, 2013) and in previous literature (Street, 2012), the presence of a z-axis offset did indeed impair the development of multi-level cognitive maps. To address this issue, we argue that the z-axis offset increases the difficulty of learning a multi-level building when the adjacent floors are overlapped, meaning that the between-floor effects found in the aforementioned studies were caused by the combination of the between-floor overlap and the z-axis offset factors rather than the z-axis offset alone. This assertion is examined in Experiment 2.

The lack of reliable differences between the 90° vs. 0° between-floor heading shift is inconsistent with the assumption based on the framework of multi-level cognitive map development. In Section 2.2.1, we predicated that when a multi-level building has no between-floor visual access and no visual access between OI-spaces, as is the case in Experiment 1, navigators have to obtain rotational displacements by updating heading information during vertical transition in order to integrate cross-level spatial knowledge. In this case, the 90° between-floor heading shift was a priori predicted to result in greater between-floor or between-region pointing errors. However, the results of Experiment 1 demonstrated that the 90° between-floor heading shift did not affect users' between-floor or between-region pointing performance. On the basis of this finding, we extended and refined the framework of multi-level cognitive map development (discussed in Section 2.2) by postulating that complex between-floor heading shifts only impair users' development of multi-level cognitive maps if navigators depend solely on path integration for cross-level spatial knowledge integration. In Experiment 1, even though all buildings had no between-floor visual access and no visual access between OI-spaces, multiple floors of these buildings were non-overlapped and shared a common reference direction. Thus, they could use interior features such as walls or hallways to learn this reference direction. In this case, navigators did not rely only on path integration for consolidating cross-level spatial knowledge and thus the 90° between-floor heading shift did not impair users' development of multi-level cognitive maps as much as anticipated. This assertion will be further studied in Experiment 3, in which we investigated whether a confusing heading shift (the combined factors of between-floor heading shift and portal-floor heading shift) would impair users' development of multi-level cognitive maps in overlapped and aligned buildings.

4.3 Experiment 2 (between-floor overlap and misalignment)

4.3.1 Introduction

The first research question addressed by Experiment 2 asks how the combination of the between-floor overlap and the z-axis offset factors affects the development of multi-level cognitive maps. In Experiment 1, we found that the z-axis offset cannot be solely attributable to the difficulty of integrating cross-level spatial knowledge. However, in the two preliminary studies (Giudice & Li, 2012; Li & Giudice, 2013) and in the work of Street (2012), the presence of a z-axis offset indeed impaired the development of multi-level cognitive maps. We argue that the between-floor effects found in the aforementioned studies were caused by the combination of the between-floor overlap and the z-axis offset factors rather than the z-axis offset alone. This assertion is examined in the current study.

Previous studies on qualitative spatial reasoning have found that direction relation between points can be implied by the relation of ancestor regions (i.e., regions the points located in) (Papadias & Egenhofer, 1997). With respect to multi-level built environments, the implication is that navigators can use the relation of two floors for the directional judgment of two positions located on the two floors. For instance, if two floors (A and B) are non-overlapped and floor A is located at the north of floor B, navigators can roughly estimate that the direction relations between two positions (*p1* on floor A and *p2* on floor B) could be north, northeast or northwest. However, if two regions are overlapped, there is information loss as no conclusion about the direction relation between points can be drawn based on ancestor regions (Papadias & Egenhofer, 1997). Thus, if floors A and floor B are overlapped, the direction relations between *p1* and *p2* could be arbitrary. In this case, when two floors of a building are overlapped, navigators cannot use their relation to imply the direction of two positions located on the two floors. In addition, as discussed in Section 4.1, when two floors of a building are overlapped, there must be a set of positions within the two floors co-located at the same x-y coordinates. If the building has no visual access between OI-spaces and no between-floor visual access, there are often no information sources to link the vertical floors/regions. Thus, the combination of the between-floor overlap and the z-axis offset factors is anticipated to increase the difficulty of integrating cross-level spatial knowledge. This assertion is investigated in Experiment 2. On the other hand, the above analysis also indicates the need for improved visualization techniques, which are evaluated in Chapter 5.

The second research question addressed by Experiment 2 asks how the floor misalignment affects the development of multi-level cognitive maps. Previous literature has found that navigators typically assume that the organization of a given floor extends to all floors (Carlson et al., 2010; Hölscher, Brösamle, & Vrachliotis, 2012). However, if two floors of a building are misaligned, this assumption is violated and navigators have to integrate the two misaligned local reference frames into a global multi-level cognitive map for accurate pointing and wayfinding between floors, which has been shown to be challenging and error-prone for humans to do accurately (see Section 2.2.1). Thus, we predicted that the presence of between-floor misalignment would impair the development of multi-level cognitive maps. This assertion is investigated in the current study. Werner and Schindler (2004) studied the effect of misalignment of local reference frames on cognitive map development in a single-floor virtual building. They systematically manipulated the orientation of an elevator, either misaligning its axis or aligning it with respect to the floor's local reference frame. The results showed that participants' pointing accuracy and wayfinding performance was significantly diminished in the misaligned condition relative to the aligned condition. However, no empirical studies have examined the effect of the between-floor misalignment on the integration of cross-level spatial knowledge, as is the focus of the current study. In this thesis, between-floor alignment information is proposed to be an important component of multi-level cognitive maps. Thus, if an effect of the between-floor misalignment is observed on users' cross-level spatial behavior performance, the findings will help validate the proposed framework of multi-level cognitive map development.

4.3.2 Methods

Sixteen participants (eight females and eight males, mean age = 21.6, *SD* = 1.8) were recruited from the University of Maine student body. All participants differed from the participants in Experiment 1.

We used the same software package and experimental apparatus as in Experiment 1 and designed four virtual buildings, seen in Figure 4.7. Each floor of the four buildings had the same floor layout as in Experiment 1. However, the four buildings were systematically manipulated based on two between-floor topological and structural properties (between-floor overlap and misalignment). In Experiment 2, buildings 3 and 4 consisted of two overlapped floors, whereas buildings 1 and 2 had only a single floor with two regions. The two regions of buildings 1 and 2 were matched with the two floors of buildings 3 and 4 in regard to layout complexity. Thus, by comparing users' performance between non-overlapped single-floor buildings (1 and 2) and overlapped two-floor buildings (3 and 4), we can examine the effect of the between-floor overlap (combined with the z-axis factor) on the development of multi-level cognitive maps. In addition, the second region/floor of buildings 2 and 4 were rotated 45° with respect to the first region/floor. By comparing users' performance between the two types of buildings (45° perspective shift vs. no perspective shift), we can examine the effect of the floor misalignment on the development of multi-level cognitive maps.

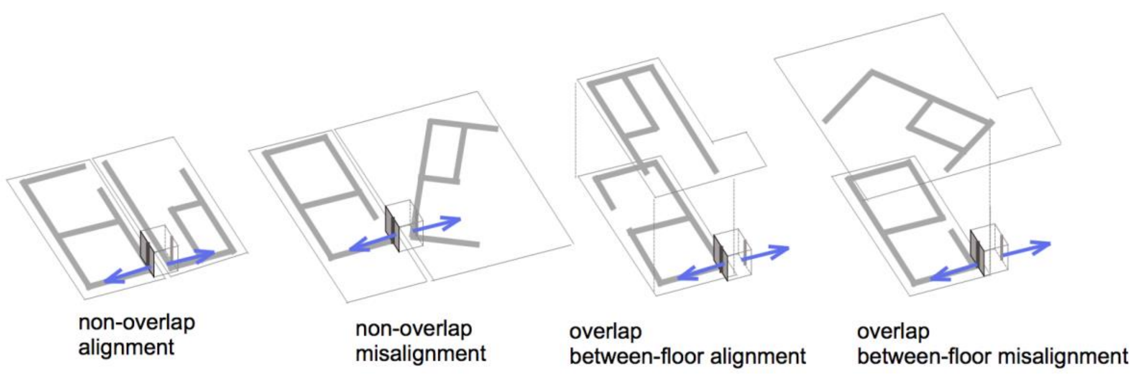


Figure 4.7. Floor layouts of Experiment 2.

The experimental procedure was the same as Experiment 1, except that the tested environments differed. All participants followed the same procedure in which they first learned a multi-level building and then took part in two cross-level spatial tasks: between-floor pointing and wayfinding (see Sections 2.4 and 3.3).

4.3.3 Results

A repeated measures ANOVA was run for each of the four dependent measures (absolute pointing error, pointing latency, wayfinding accuracy, and wayfinding efficiency) with two within-subject factors (between-floor overlap and misalignment). Significant main effects of between-floor overlap were observed for both pointing error and pointing time, with pointing in non-overlapped buildings being reliably faster and more accurate than in overlapped buildings: pointing error, $F(1, 63) = 56.977$, $p < .0001$, $\eta^2 = .475$; and pointing time, $F(1, 63) = 10.215$, $p < .005$, $\eta^2 = 140$. Significant main effects of misalignment were observed for all measures: pointing error, $F(1, 63) = 37.366$, $p < .0001$, $\eta^2 = .372$; pointing time, $F(1, 63) = 14.856$, $p < .0005$, $\eta^2 = 191$; wayfinding accuracy, $F(1, 31) = 5.423$, $p < .05$, $\eta^2 = .149$; and wayfinding efficiency, $F(1, 31) = 6.249$, $p < .05$, $\eta^2 = .168$.

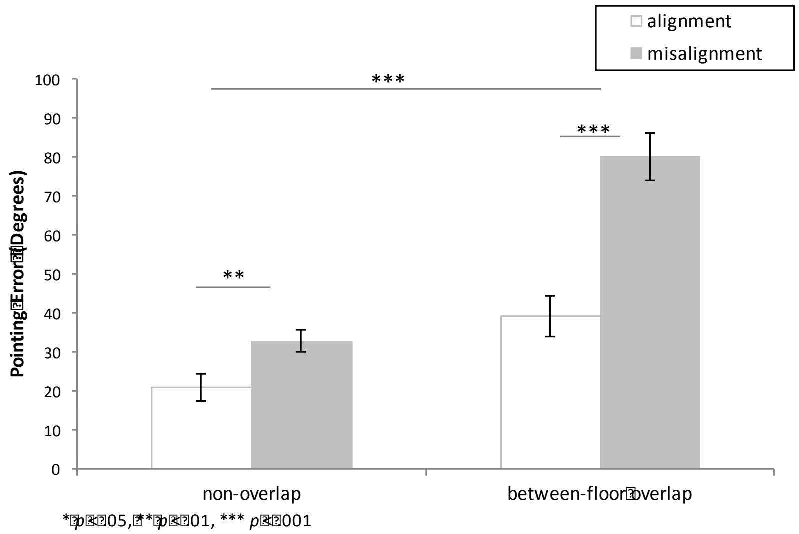


Figure 4.8. Mean absolute pointing error for Experiment 2.

The interaction effect between misalignment and between-floor overlap was significant for pointing error and pointing time: pointing error, $F(1, 63) = 12.226$, $p < .001$, $\eta^2 = .163$, and pointing time, $F(1, 63) = 4.920$, $p < .05$, $\eta^2 = .072$. Subsequent Dunn–Sidak pairwise comparisons showed that both interactions were driven by the between-floor misalignment condition, which took longer and had a larger error than the two alignment conditions (all ps < .001).

Table 4.2. Mean absolute pointing error, pointing latency, wayfinding accuracy, and wayfinding efficiency in Experiment 2.

Dependent Variables	Independent Variable II	Independent Variable I	
		Non-overlap	Between-floor overlap
Absolute pointing error	Alignment	20.84 (3.43)	38.95 (5.17)
	Misalignment	32.62 (2.86)	80.04 (6.15)
Pointing latency	Alignment	14.70 (1.18)	17.07 (1.42)
	Misalignment	19.80 (1.76)	30.88 (4.15)
Wayfinding accuracy	Alignment	87.5% (5.9%)	87.5% (5.9%)
	Misalignment	75.0% (7.8%)	71.9% (8.1%)
Wayfinding efficiency	Alignment	85.1% (5.9%)	87.5% (5.9%)
	Misalignment	73.3% (7.6%)	71.1% (8.0%)

4.3.4 Discussion

The primary goal of Experiment 2 was to investigate whether the combination of the between-floor overlap and the z-axis offset factors or the floor misalignment factor impairs the development of multi-level cognitive maps. The results demonstrated that participants took significantly more time and made larger pointing errors in the overlapped conditions than in non-overlapped conditions, suggesting that between-floor overlap undermined users' development of multi-level cognitive maps. However, no effects of the between-floor overlap were observed for the two wayfinding measures. Our lack of an effect may be due to our environments, as the two floors of all Experiment 2 buildings were connected by only one elevator, so the between-floor overlap factor did not affect the between-floor connectivity information. According to the logical relation of multi-level cognitive maps and cross-level spatial tasks (see Chapter 2.4), the between-floor overlap factor should not affect the between-floor wayfinding performance. Thus, the finding of the between-floor overlap is consistent with the predictions based on the framework of multi-level cognitive map development. The effect of the expected between-floor overlap on between-floor pointing performance was observed, suggesting that it is more difficult to maintain the spatial relation of objects between overlapped floors than non-overlapped environments. Thus, the assertion that the between-floor effects observed in the two preliminary studies and in the one by Street (2012) were caused by the combination of the between-floor overlap and the z-axis offset rather than the z-axis offset alone is validated. In the real world, almost all buildings are designed with fully or partially overlapped floors, meaning that the between-floor overlap property is one of the most prominent topological characteristics of multi-floor buildings. Results of Experiment 2 showed that there is a trade-off between the benefits of efficient use of land space and the increased difficulty in forming a multi-level cognitive map. This issue will be addressed by improved visualization interfaces (see Chapter 5).

The results of Experiment 2 revealed that participants had greater pointing errors, longer pointing latencies, and lower wayfinding accuracy and efficiency in misaligned buildings than in aligned buildings, providing compelling evidence that the floor misalignment is a substantial factor leading to the difficulty of developing accurate multi-level cognitive maps. This finding is consistent with the predictions based on the framework of multi-level cognitive maps (Sections 2.2-2.4). Given that in the current study two floors of a building were misaligned, the misalignment factor violated the assumption that two floors of a building share a common spatial reference frame. As a result, navigators had to integrate misaligned cross-level spatial knowledge into a multi-level cognitive map based on their path integration (all Experiment 2 buildings had no visual access between OI-spaces and no between-floor visual access). Thus, users' between-floor pointing performance was significantly diminished in the misaligned conditions relative to the aligned conditions. In addition, the misalignment factor causes a portal-floor heading shift on the second floor (β2), meaning that the between-floor connectivity information was affected. Thus, users' between-floor wayfinding performance was also undermined in the misaligned conditions relative to the aligned conditions. These findings offer an important implication for architectural design, namely, when a multi-level building has no between-floor visual access and no visual access between OI-spaces, misaligned floors should be avoided and a common spatial reference frame between floors is critically important for navigators to be able to integrate cross-floor spatial knowledge. For instance, the common spatial reference frame can be created by making the salient common axes or boundaries common between floors. However, in the physical world, a common spatial reference frame is not always available and making the salient axes common through structural modifications is impractical. Thus, improved visualization interfaces, as are investigated in Chapter 5, play an important role in assisting users to learn this important information regarding between-floor misalignment.

4.4 Experiment 3 (45° portal-floor heading shift)

4.4.1 Introduction

In Experiment 1, we found that the 90° between-floor heading shift does not impair the development of multi-level cognitive maps in non-overlapped buildings. On this basis, we argued that complex between-floor heading shifts only impair users' development of multi-level cognitive maps if navigators solely depend on path integration for cross-level spatial knowledge integration. This assertion is further investigated in the current study. In Experiment 3, we designed two overlapped and aligned buildings using VEs, as pictured in Figure 4.9. The design was similar to that of Experiment 2, except that the elevator had both a 90° between-floor heading shift and a 45° portal-floor angular offset, depicted in Figure 4.9.

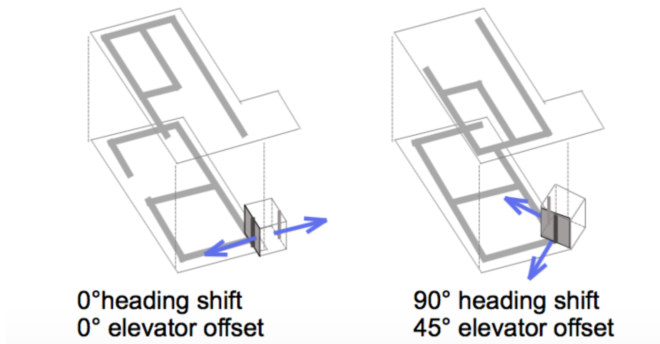


Figure 4.9. Floor layouts of Experiment 3.

In Experiment 3, we investigated whether a confusing heading shift (the combined factor of between-floor heading shift α and portal-floor heading shift β, called misaligned portals) would impair users' development of multi-level cognitive maps. All buildings of Experiment 3 lacked between-floor visual access and no visual access between OI-spaces. However, all buildings were overlapped and vertically-aligned, meaning that they all shared a common spatial reference frame. Based on the logical relation between multi-level cognitive maps and cross-level spatial tasks, we postulated that the confusing heading shift after vertical transition would increase the difficulty of path integration and subsequently impair users' development of multi-level cognitive maps. However, as described earlier, if navigators could use interior features such as walls or hallways to learn the common spatial reference frame, they would not need to depend on path integration for cross-level spatial knowledge integration. Thus, the combined factors of between-floor heading shift and portal-floor heading shift would not impair users' development of multi-level cognitive maps.

4.4.2 Methods

Sixteen new participants (eight females and eight males, mean age = 20.2, *SD* = 1.2) were recruited from the University of Maine student body. We used the same software package and experimental apparatus as in Experiment 2. The experimental procedure was the same as Experiment 1, except that different experimental environments were used.

4.4.3 Results

A one-way repeated measures ANOVA was run for each of the four dependent measures (absolute pointing error, pointing latency, wayfinding accuracy, and wayfinding efficiency), with one within-subject factor of misaligned portals (the combined factor of between-floor heading shift and portal-floor heading shift). Although a significant main effect was observed for pointing latency, $F(1, 63) = 4.56$, $p < .05$, $\eta^2 = .068$, no effect was found for any of the other three measures.

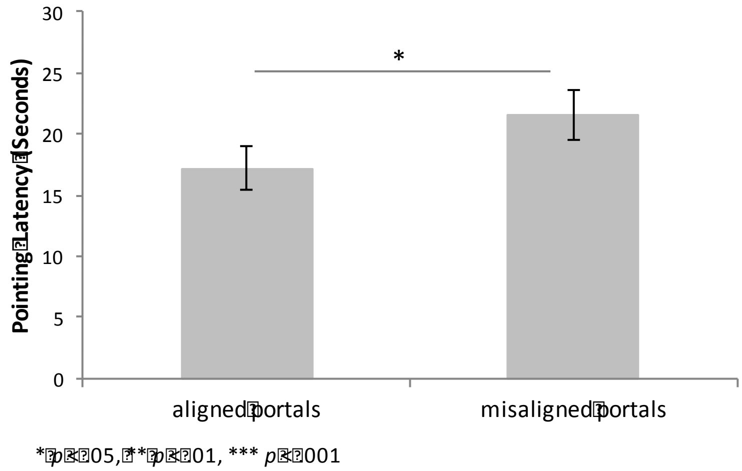


Figure 4.10. Mean pointing latency for Experiment 3.

Based on the observed effect sizes, we used G*power to determine how many subjects would have been necessary for statistical significance at power level of .95. We found that in order to find statistical significance, sample size was not dramatically increased (n < 1,000), suggesting that the effect of misaligned portals might exist but it was not observed in this study. However, the bottom line is the overall results indicate that the misaligned portals factor did not reliably impair users' development of multi-level cognitive maps in aligned buildings.

Table 4.3. Mean absolute pointing error, pointing latency, wayfinding accuracy, and wayfinding efficiency in Experiment 3.

Dependent Variables	Independent Variable I	
(Measures)	Aligned portals	Misaligned portals
Absolute pointing error	32.22 (3.38)	28.67 (3.24)
Pointing latency	17.17 (1.75)	21.51 (2.04)
Wayfinding accuracy	90.6% (5.2%)	78.1% (7.4%)
Wayfinding efficiency	89.0% (5.2%)	75.9% (7.3%)

4.4.4 Discussion

The primary goal of Experiment 3 was to investigate whether the combined factors of between-floor heading shift and portal-floor heading shift (misaligned portals) would impair users' development of multi-level cognitive maps in an aligned building. The results showed that the combined factor (misaligned portals) caused only slightly longer pointing latency, and no reliable differences of absolute pointing error, wayfinding accuracy and wayfinding efficiency were observed between the control condition (no heading shift) and the misaligned portals condition (confusing heading shift). This lack of an effect on the absolute pointing error is likely due to the environments tested, as all Experiment 3 buildings were aligned and navigators might have used interior features such as hallways or walls to learn the common spatial reference frame between floors. It should be noted that in the learning phase navigators needed to learn the building twice and then took part in a criterion pointing task to ensure that they had built accurate single-level survey knowledge. During the two phases, navigators might have deduced that the two floors of Experiment 3 buildings were aligned and shared a common spatial reference. Thus, the confusing heading shift (the combined factor of between-floor heading shift and portal-floor heading shift) surprisingly did not largely impair users' ability to form multi-level cognitive maps. This finding validates our assertion that complex between-floor heading shifts only impair users' development of multi-level cognitive maps if navigators solely depend on the path integration process for cross-level spatial knowledge integration. Nevertheless, increased pointing latency suggests that there was additional cognitive effort required to perform the between-floor pointing task and it is unclear whether the confusing heading shift leads to the difficulty of learning a building with misaligned floors. These issues will be addressed in future studies.

4.5 Experiment 4 (global landmarks)

4.5.1 Introduction

In Section 2.2.2, we argued that visual access to a global landmark from within a building promotes users' development of multi-level cognitive maps (*Hypothesis 2*), and if a global landmark is visible from both indoor and outdoor spaces, it will facilitate both cross-level spatial knowledge integration and the integration of OI-spaces (*Hypothesis 3*). To evaluate these two hypotheses, Experiment 4 examined the effects of global landmarks (both indoor and outdoor) on users' development of multi-level cognitive maps.

Global landmarks are salient environmental features visible at a large spatial scale from within the environment. Previous literature on outdoor wayfinding has found clear evidence that these global landmarks provide a fixed spatial reference frame for navigators to integrate local spatial knowledge into a global cognitive map (see Steck & Mallot (2000) for review). However, there is no empirical evidence on the effect of global landmarks observed from within a building in supporting users' ability to form multi-level cognitive maps. In a previous study, we investigated whether two vertically-aligned chandeliers co-located on separate floors, called contiguous landmarks, could serve as a global landmark and facilitate users' development of a multi-level cognitive map (see Experiment 7). However, we observed no reliable effects of contiguous indoor landmarks and very few users even noticed that the chandeliers were vertically aligned. We interpreted the absence of an effect as owing to the fact that users had to perceive each chandelier discretely on separate floors, making it hard for them to mentally link the two inter-floor locations without having direct access to each other.

These results suggest that indoor global landmarks for multi-level built environments need to be more than co-located at the same x-y coordinates between floors, they must also be directly perceivable from multiple locations/levels of the building. Therefore, in Experiment 4 we designed an outdoor global landmark (a church) and an indoor global landmark (a statue in an atrium), both of which were visible from within the building over multiple locations, as shown in Figure 4.11.

Figure 4.11. Outdoor and indoor global landmarks.

A global landmark, serving as a fixed global spatial reference, helps users consolidate single-level spatial knowledge into a consistent/global multi-level cognitive map. For instance, navigators can use global landmarks to learn the between-floor perspective shift γ and to estimate relative direction between two positions located on two floors (see Section 2.2.2 for details). Thus, we propose that users would develop a more accurate multi-level cognitive map when they could see the global landmark from both floors rather than only from a single floor. We predict that both a statue in an atrium and an external landmark can serve as a global landmark, as they are directly perceivable from multiple locations/levels of a building. This assertion was evaluated in the current study.

The present research also aims to investigate whether visual access to global landmarks can facilitate users' integration of outdoor and indoor spaces, which has attracted increasing attention in recent years (see (Giudice et al., 2010) for review). In the current studies, OI-space integration was measured by pointing latency and error performance when pointing from indoor locations (e.g., the building's rooms) to an outdoor location, e.g., a parking lot.

4.5.2 Methods

Sixteen participants (eight females and eight males, *M*=20.1, *SD*=2.0) were recruited from the University of Maine student body.

The experimental environment was displayed on a Samsung 43" Class Plasma HDTV monitor running at 60 Hz and at a resolution of 1024 × 768. The desktop VEs were run with a MacBook Pro (2.2 GHz Intel Core i7). The Unity 5.1 VR engine (Unity Technologies) was used as the VE platform supporting users' real-time navigation and recording their trajectory and test performance. Our environments comprised four two-level buildings, as shown in Figure 4.12. Participants used an elevator to move between floors. All buildings were matched for layout complexity and topology.

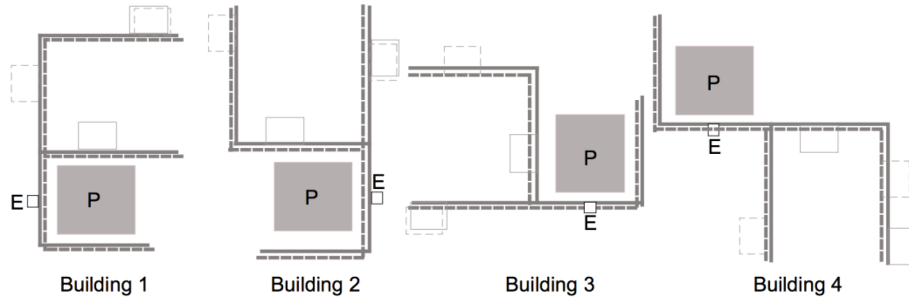


Figure 4.12. Floor layouts of Experiment 4. The solid line represents the first-floor layout and the dashed line represents the second-floor layout.

Each virtual building contained four target rooms: a bathroom, a dining room, a conference room, and an office. In addition, each environment had a number of empty rooms evenly located in the building, as shown in Figure 4.13. A set of fire extinguishers or water fountains were located directly above/below target rooms, and served as the targets for the drilling task, as described in the experimental procedure. Each environment included a global landmark—either a church or a statue in an atrium—visible from a single floor or from both floors. As shown in Figure 4.13, each floor included a number of windows, through which users had visual access to the global landmark. Each environment also contained a parking lot. Participants were positioned at the parking lot at the beginning of the experiment. However, when inside the building, the parking lot was only visible from the window opposite the elevator, as shown in Figure 4.13. Thus, the parking lot was not a global landmark in the current studies, but it served as a fixed geo-reference for the outdoor environment. We tested users' integration between indoor and outdoor spaces by asking them to point from rooms inside the building to this parking lot.

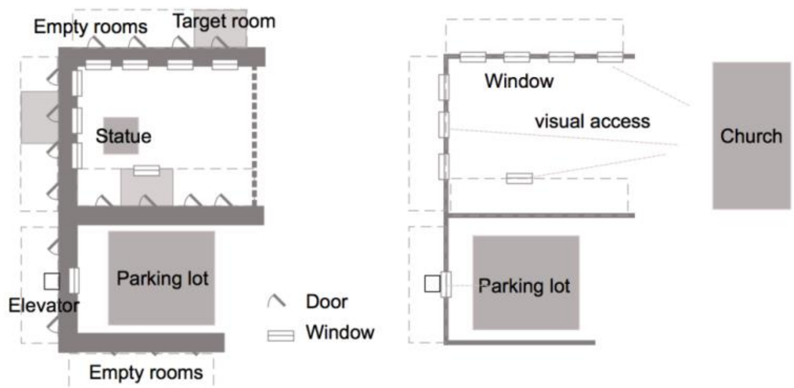


Figure 4.13. Visual access to the indoor and outdoor global landmark.

A within-subject design was adopted, with the sixteen participants running in all four conditions: (1) single-floor visual access to an outdoor global landmark, (2) single-floor visual access to an indoor global landmark, (3) two-floor visual access to an outdoor global landmark, and (4) two-floor visual access to an indoor global landmark). All participants in Experiments 4-6 followed the same procedure: they first learned a multi-level building and then took part in three cross-level spatial tasks: pointing and wayfinding between floors and a drilling task (see Section 2.4 and 3.3 for more details).

4.5.3 Results

The five dependent measures (pointing latency, absolute pointing error, wayfinding accuracy, wayfinding efficiency, and drilling accuracy) were analyzed for each participant. A 2 (visual access: single-floor vs. two-floor) × 2 (global landmark type: indoor vs. outdoor) × 3 (pointing target type: global landmark, parking lot, and building rooms) repeated-measures ANOVA was conducted for each of the two dependent measures of pointing latency and absolute pointing error. Significant main effects of visual access were observed for both measures, with pointing in the two-floor visual access condition being faster and more accurate than pointing in the single-floor visual access condition: pointing latency, $F(1, 63) = 11.151$, $p < .001$, $\eta^2 = .150$; and absolute pointing error, $F(1, 63) = 10.057$, $p < .005$, $\eta^2 = .138$.

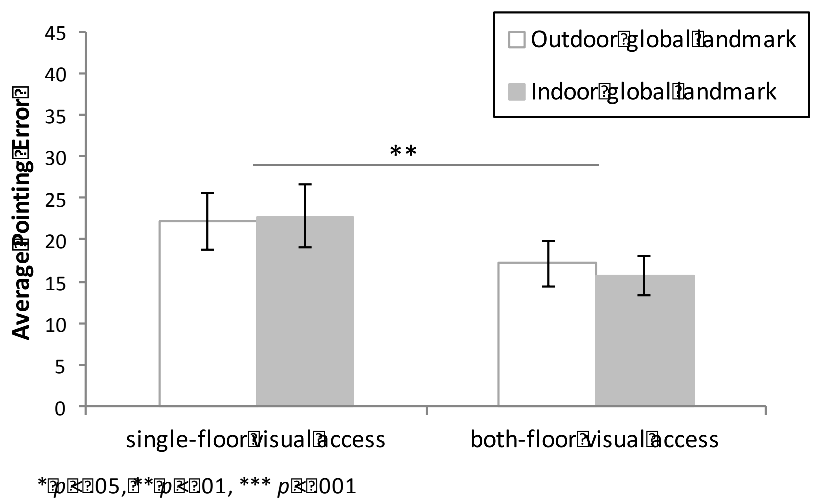


Figure 4.14. Mean absolute pointing error for Experiment 4.

Significant main effects of target type were also observed for both pointing latency and absolute pointing error: latency, $F(2, 126) = 58.361$, $p < .0001$, $\eta^2 = .481$; and error, $F(2, 126) = 15.631$, $p < .0001$, $\eta^2 = .199$. Subsequent Dunn–Sidak pairwise comparisons showed that pointing to the global landmark was faster and more accurate than pointing to the parking lot and the internal rooms (all *p*s < .001). A significant global landmark type by pointing type interaction was observed for pointing error, $F(2, 126) = 7.198$, $p < .001$, $\eta^2 = .103$. Subsequent Dunn–Sidak pairwise comparisons demonstrated that this significant interaction was driven by the trials requiring pointing to the parking lot, which was reliably more accurate in the outdoor global landmark conditions than with the indoor global landmark conditions (all *p*s < .05).

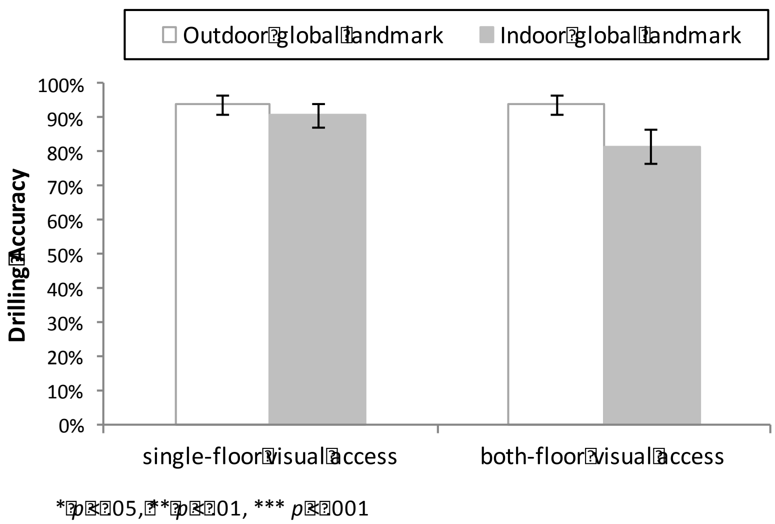


Figure 4.15. Mean drilling accuracy for Experiment 4.

A 2 (visual access) × 2 (global landmark type) repeated-measures ANOVA was conducted for each of the three dependent measures of wayfinding accuracy, wayfinding efficiency, and drilling accuracy. A significant main effect of global landmark type was observed for drilling accuracy, with drilling performance in the outdoor global landmark condition found to be more accurate than performance in the indoor global landmark condition, $F(1, 63) = 4.817$, $p < .05$, $\eta^2 = .071$. There were no significant main effects of visual access (all ps > .05) or global landmark type (all ps > .05) on wayfinding accuracy or wayfinding efficiency.

Table 4.4. Mean absolute pointing error and pointing latency in Experiment 4.

Dependent Variables	Pointing target type	Global landmark type	Visual access	
			Two-floor	Single-floor
Absolute pointing error	Global landmark	Outdoor	12.33 (1.96)	19.42 (2.80)
		Indoor	6.57 (0.91)	13.56 (3.76)
	Parking lot	Outdoor	22.63 (3.26)	19.55 (2.56)
		Indoor	25.090 (4.07)	33.31 (4.45)
	Building rooms	Outdoor	16.42 (2.85)	27.99 (4.49)
		Indoor	15.21 (1.98)	21.73 (3.25)
Pointing latency	Global landmark	Outdoor	3.77 (0.28)	7.07 (1.19)
		Indoor	4.05 (0.23)	5.02 (0.44)
	Parking lot	Outdoor	6.83 (0.69)	8.09 (0.82)
		Indoor	7.95 (0.95)	8.59 (0.90)
	Building rooms	Outdoor	12.19 (1.50)	13.10 (1.71)
		Indoor	11.22 (1.28)	14.77 (1.38)

Table 4.5. Mean wayfinding accuracy, wayfinding efficiency, and drilling accuracy in Experiment 4.

Dependent Variables	Pointing target type	Global landmark type	Visual access	
			Two-floor	Single-floor
Wayfinding		Outdoor	78.1% (5.2%)	70.3% (5.8%)
accuracy		Indoor	71.9% (5.7%)	65.6% (6.0%)
Wayfinding		Outdoor	75.3% (5.2%)	67.3% (5.7%)
efficiency		Indoor	67.1% (5.5%)	63.1% (5.9%)
Drilling		Outdoor	93.8% (3.0%)	93.8% (3.0%)
accuracy		Indoor	81.3% (4.9%)	90.6% (3.7%)

4.5.4 Discussion

The primary goal of Experiment 4 was to investigate whether increasing visual access to an indoor or outdoor global landmark observed through the building's windows would assist users' development of a multi-level cognitive map. As we predicted, the results demonstrated that users' pointing was reliably faster and more accurate in the two-floor visual access condition than in the single-floor visual access condition, providing clear evidence that increasing visual access to a global landmark (both indoor and outdoor) through direct window access significantly promoted users' ability to form multi-level cognitive maps. This finding supports our hypothesis that both an outdoor and indoor global landmark can serve as a fixed spatial reference frame for navigators to integrate multi-level spatial knowledge into a globally coherent multi-level cognitive map, which provides validation for *Hypothesis 2* of this thesis. The results also demonstrated that the outdoor global landmark not only aided with the development of multi-level cognitive maps, but also assisted with the integration of indoor and outdoor spatial reference frames, thereby providing corroborating evidence for *Hypothesis 3* of this thesis. Previous literature has discussed that increasing visual access to important level-related building features such as elevators could support users' spatial learning and wayfinding of a multi-level building (Giudice & Li, 2012; Hölscher et al., 2006). Our current research extends these earlier studies and demonstrates that increasing visual access to a global indoor or outdoor landmark can also improve the performance of between-floor pointing. This research provides important empirical evidence for the framework of multi-level cognitive map development (discussed in Section 2.2.2); for instance, users could learn between-floor alignment by computing the bearing difference to a global landmark rather than constantly updating their heading directions during vertical travel. In this case, the difficulty of learning a multi-level building with confusing elevators/staircases could be greatly reduced (or alleviated) if visualization interfaces can assist navigators to have access to a global landmark (see Chapter 5). Experiment 4 results provide important empirical foundations for the design of Augmented Reality (AR) models used in Experiments 5 and 6, which aim to use AR technology to extend the benefit of global landmarks providing a fixed spatial reference frame to buildings that otherwise do not have visual access to this cue.

There was a small reversal of the effect of global landmark type on drilling accuracy, suggesting that the outdoor global landmark was more efficient for promoting users' learning of vertical alignment information than the indoor global landmark. However, the predicted effect of visual access on drilling accuracy was not observed, meaning that increased visual access to the landmark from both floors did not help users learn accurate between-floor alignment information. We believe that drilling accuracy may have been elevated in Experiment 4 because the fire extinguishers and water fountains were always located directly above/below a target room and participants could have used this as a cue. This issue is addressed in Experiment 5.

4.5 Summary

In Chapter 4, we described four experiments (Experiments 1-4) using VEs to evaluate (1) five multi-level structural and topological properties, and (2) global indoor and outdoor landmarks, on the development of multi-level cognitive maps. Experiments 1-3 were conducted to evaluate *Hypothesis 1*, which asserts that between-floor structural and topological properties of buildings impair users' development of multi-level cognitive maps. Hypothesis 1 was partially validated. Results of Experiment 2 demonstrated that the between-floor overlap and the misalignment factors substantially increased the difficulty of building multi-level cognitive maps. Results of Experiments 1 and 3, however, showed that the z-axis offset, the between-floor heading shift and the combined factor (the between-floor heading shift and the portal-floor heading shift) did not make it more difficult to form multi-level cognitive maps in aligned buildings (overlapped or non-overlapped). The findings of Experiments 1-3 were generally consistent with the predictions based on the framework of multi-level cognitive map development (Section 2.2). However, on the basis of the findings of Experiment 1 (based on the 90° between-floor heading shift factor), we extended the framework of multi-level cognitive map development to reflect that complex between-floor heading shifts only impair users' development of multi-level cognitive maps if navigators solely depend on path integration for cross-level spatial knowledge integration. This assertion was further validated in Experiment 3. Experiment 4 was conducted to evaluate *Hypothesis 2*, which asserts that visual access to a global landmark can promote users' development of multi-level cognitive maps) and *Hypothesis 3*, which asserts that if a global landmark is visible from both indoor and outdoor spaces, it will facilitate both cross-level spatial knowledge integration and the integration of OI-spaces. Both of these hypotheses were validated by the results of Experiment 4, providing important empirical foundations for the design of the visualizations used in Experiments 5 and 6, which aim to use AR technology to extend the benefit of global landmarks to buildings that otherwise do not have visual access to this cue.

CHAPTER 5

EVALUATION OF VISUALIZATION INTERFACES FOR ASSISTING THE DEVELOPMENT OF MULTI-LEVEL COGNITIVE MAPS

5.1 Introduction

In Section 2.2.3, we discussed that global landmarks are often not available in multi-level indoor environments, so the advantage of these global landmarks—serving as a fixed spatial reference frame—is often greatly reduced when learning and navigating within buildings (Giudice et al., 2010). However, we argued that we can use augmented reality (AR) technology to increase visual access to global landmarks, which could facilitate users' development of multi-level cognitive maps (*Hypothesis* 4). In this chapter, we conducted two experiments (Experiments 5-6) using VEs to evaluate this hypothesis. In addition, Experiment 7 was conducted using VEs to evaluate *Hypothesis 5:* schematic maps that effectively convey the desired multi-level building information to users could alleviate the challenge of integrating cross-level spatial knowledge.

5.2 Experiment 5 (icon-model and wireframe-model)

5.2.1 Introduction

The results of Experiment 4 showed that increasing visual access to a global landmark observed through the building's windows promoted users' development of multi-level cognitive maps. However, as discussed earlier, direct access to global landmarks is often not available from within buildings and increasing visual access through structural modifications is impractical. Thus, the current study (Experiment 5) aimed to use AR technology to extend the benefits found in Experiment 4 to many buildings without physical visual access to global landmarks. We proposed and evaluated two AR models to improve visualization (an icon-model vs. a wireframe-model), as shown in Figure 5.1.

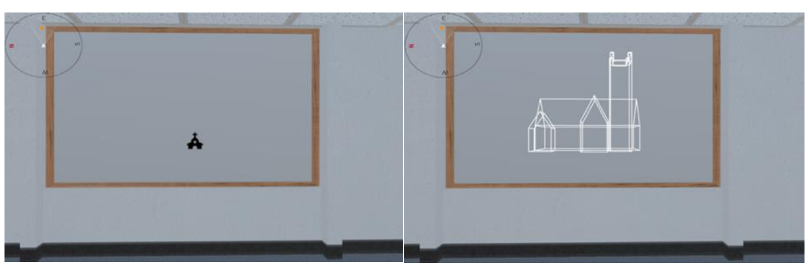


Figure 5.1. Icon-model (left) and wireframe-model (right) of the global landmark.

An icon-model uses a visual symbol to indicate the global landmark's direction. By contrast, a wireframe-model indicates not only the direction of the global landmark, as the icon-model does, but also the perspective from which users can see the landmark, and its edges, as shown in Figure 5.1. Users' performance with the two AR visualization techniques were compared to two control conditions: (1) no visual access to outdoor spaces, which is the baseline control condition, and (2) a window-access condition. These two AR models require fewer computational resources to render and take less time to create when compared to other visualization techniques, as reviewed in (Dey & Sandor, 2014). Thus, if one (or both) were found to be as efficient as the window-access condition in facilitating multi-level cognitive map development and subsequent cross-floor spatial behaviors, we would have an economical and broad-based solution for improving indoor visualization.

5.2.2 Methods

Sixteen unique students participated in Experiment 5. The design was similar to that of Experiment 4, except for the following changes. First, only the church was used as the global landmark. Second, the locations of the fire extinguishers and water fountains were adjusted to ensure that only a subset of them were vertically aligned with a target room.

5.2.3 Results

A repeated-measures ANOVA was conducted for each of the two dependent measures of pointing latency and absolute pointing error, with the four conditions of visual access and three pointing target types as two within-subject factors. A significant main effect of visual access was observed for absolute pointing error, $F(3, 189) = 14.925$, $p < .0001$, $\eta^2 = .192$, with pointing in the window-access condition being more accurate than the no visual access condition and the two AR interface conditions (all ps < .0001). This finding suggests that the visualization of the global landmark provided by the two AR conditions was not as effective as the "gold standard" of direct window access in assisting users' development of a multi-level cognitive map. Significant main effects of target type on pointing performance were observed for both pointing latency and absolute pointing error: latency, $F(2, 126) = 25.420$, $p < .0001$, $\eta^2 = .287$; and error, $F(2, 126) = 7.175$, $p < .001$, $\eta^2 = .102$. Subsequent Dunn–Sidak pairwise comparisons showed that pointing performance to the global landmark was more accurate than pointing to the parking lot ($p < .005$) but not more accurate than pointing to the building's rooms ($p = .080$). Even though users were assisted with the AR visualization of the global landmark (i.e., the church), no reliable differences were found between pointing to the church and to the building's rooms, suggesting that the two AR models were not as effective as direct window access in enhancing users' spatial awareness of the church and thus, it failed to serve as a "global landmark" in this study. One explanation for this result is the lack of depth information about the global landmark within the two AR models. Without this depth information, users may have perceived the global landmark to be "floating" in space, leading to an erroneous perception of its true location. In addition, no outside boundary information of the building was visible from the AR visualizations, as could be seen through the building's windows.

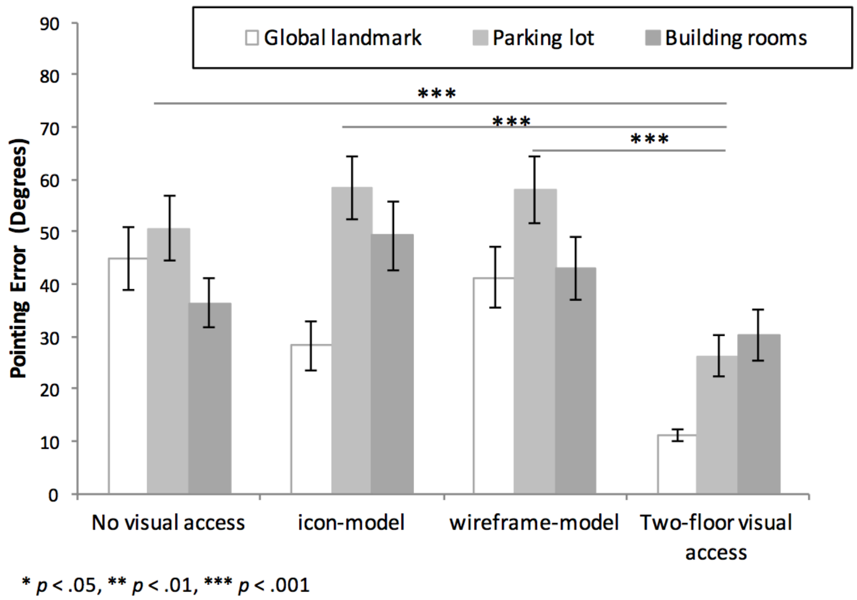


Figure 5.2. Mean absolute pointing error for Experiment 5.

A repeated-measures ANOVA was conducted for each of the three dependent measures of wayfinding accuracy, wayfinding efficiency, and drilling accuracy, with the four conditions of visual access as a within-subject factor. There was no significant main effect of visual access for any measure (all *p*s > .05). The average drilling accuracy (M = 57.4%, SE = 1.9%) was significantly lower than that of Experiment 4 (M = 89.8%, SE = 1.9%), $t(510)=8.935$, $p < .0001$, supporting our assertion that the design of the buildings in Experiment 4 artificially elevated users' drilling accuracy performance. Even with these modifications, drilling accuracy was still not promoted by the window-access condition, suggesting that direct visual access to a global landmark alone does not facilitate users' learning of between-floor alignment. The drilling task requires accurate between-floor alignment information, which was not sufficiently provided by global landmarks in the current study. We believe that to promote drilling accuracy, the AR interface must also assist users to visualize the objects above/below their current location. This assertion is evaluated in Experiment 6.

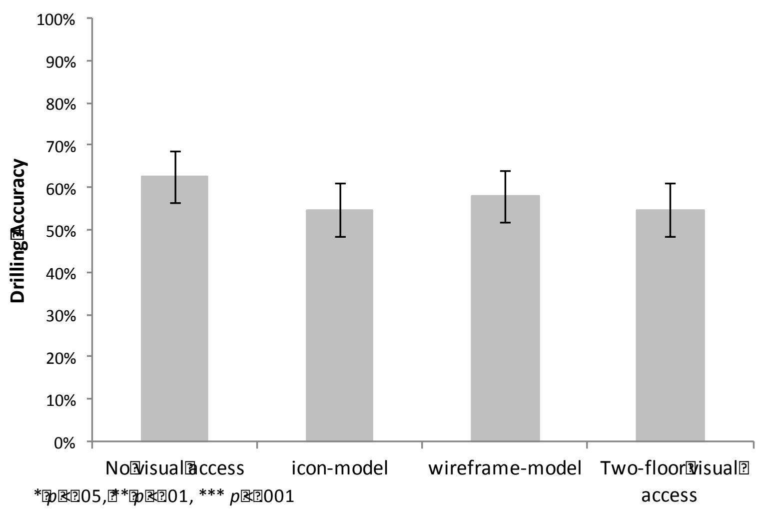


Figure 5.3. Mean drilling accuracy for Experiment 5.

Table 5.1. Mean absolute pointing error, pointing latency, wayfinding accuracy, wayfinding efficiency, and drilling accuracy in Experiment 5.

Dependent Variables	Pointing target type	Visual access			
		No visual access	Icon-model	Wireframe-model	Two-floor visual access
Absolute pointing error	Global landmark	44.89 (5.95)	28.18 (4.84)	41.22 (5.92)	11.07 (1.22)
	Parking lot	50.57 (6.15)	58.34 (6.16)	58.01 (6.43)	26.26 (4.06)
	Building rooms	36.37 (4.61)	49.26 (6.53)	42.88 (6.07)	30.22 (5.00)
Pointing latency	Global landmark	8.80 (1.40)	6.02 (0.72)	7.76 (1.33)	4.10 (0.35)
	Parking lot	7.00 (1.13)	6.19 (0.789)	7.18 (0.90)	7.12 (1.03)
	Building rooms	14.08 (2.15)	13.82 (1.83)	13.28 (2.17)	11.99 (1.48)
Wayfinding accuracy		51.6% (6.3%)	40.6% (6.2%)	40.6% (6.2%)	46.9% (6.3%)
Wayfinding efficiency		50.9% (6.3%)	40.0% (6.1%)	40.0% (6.1%)	46.0% (7.2%)
Drilling accuracy		62.5%(6.1%)	54.7% (6.3%)	57.8% (6.2%)	54.7% (6.3%)

5.2.4 Discussion

The primary goal of Experiment 5 was to investigate whether increasing visual access to a global landmark using two AR interfaces (an icon-model and a wireframe-model) would assist users' development of a multi-level cognitive map. The results of Experiment 5 showed that the two simply rendered AR models, although resource efficient, did not provide sufficient visualization fidelity, and thus, were not effective for facilitating multi-level cognitive map development.

5.3 Experiment 6 (X-ray AR visualization)

5.3.1 Introduction

The AR visualization models in Experiment 5 had three shortcomings: (1) they provided no depth information about the global landmark, (2) they could not help users perceive what was directly above or below their current location, and (3) users were constantly exposed to the AR information through an always-on interface. On the basis of the Experiment 5 findings and acknowledging these limitations, we redesigned an X-ray visualization in Experiment 6 by allowing navigators to see transparent walls, the global landmark, and the horizon of the outdoor space, as shown in Figure 5.4. The X-ray visualization provided access to depth information about the global landmark, similar to the access afforded through the building's windows. Thus, it is anticipated to be as efficient as direct window access in assisting users' development of multi-level cognitive maps. Importantly, the X-ray visualization also facilitates users to perceive what is directly above or below their current location. Thus, it is also predicted that users' drilling accuracy will be promoted by access to this AR interface in Experiment 6. In addition, users could turn on/off the AR information on-demand.

Figure 5.4. An X-ray visualization with depth information.

A second goal of Experiment 6 was to investigate whether visual access to multiple global landmarks is more efficient than visual access to a single global landmark for users' development of multi-level cognitive maps. Previous literature has discussed several methods for how humans use landmarks for self-localization, such as computing position using bearing and distance to a single landmark, computing position using distances to multiple landmarks (trilateration), and computing position using bearings or bearing differences to multiple landmarks (triangulation), as reviewed in (Loomis, et al., 1999). Visual access to multiple global landmarks has been used to help self-localization in outdoor spaces (see (Steck & Mallot, 2000) for review). However, little is known about the effect of having visual access to multiple global landmarks in multi-level built environments and there is no empirical evidence on the effect of access to global landmarks perceived through AR interfaces on users' development of a multi-level cognitive map. This issue is evaluated in the current study.

In Experiment 6, we evaluated the X-ray visualization with two global landmark conditions (single global landmark access vs. multiple global landmark access), compared to two control conditions (no visual access to outdoor spaces vs. direct window-access), as were used in Experiment 5. In addition to the church, four distinctive town houses were located on one side of the building, serving as landmarks. In the single global landmark access condition, only the church was visible through the X-ray visualization, whereas in the multiple global landmarks condition, both the houses and the church were visible throughout the building via the X-ray visualization.

5.3.2 Methods

Sixteen unique students participated in Experiment 6. The design was similar to that of Experiment 5, except that there was only one visualization interface evaluated but with two global landmark conditions.

5.3.3 Results

A repeated-measures ANOVA was conducted for each of the two dependent measures of pointing latency and absolute pointing error, with the four conditions of visual access and three pointing target types as two within-subject factors. A significant main effect of visual access was observed for absolute pointing error, $F(3, 189) = 10.746$, $p < .0001$, $\eta^2 = .146$, with pointing in the X-ray visualization (single global landmark access) condition and the window-access condition being more accurate than the no visual access condition (all ps < .0005). We found compelling evidence that the X-ray visualization (single global landmark access) is as effective as the gold standard of window-access in promoting users' development of multi-level cognitive maps. Note that, users' pointing error to the parting lot was even numerically larger in the window-access condition than in the X-ray visualization condition. With the assistance of the X-ray visualization, users had visual access to the global landmark, the parking lot and the building's rooms from anywhere in the building. Thus, they could learn the spatial relations between places within the multi-level built environment from any location, which was not the case with windows-based access, and this increased spatial visualization aided the development of a multi-level cognitive map.

No significant effect between the two global landmark conditions of the X-ray visualization was observed (single global landmark access vs. multiple global landmark access) (p > .05). This result suggests that increasing visual access to multiple global landmarks did not improve multi-level cognitive mapping performance. The larger numeric absolute pointing error observed in the multiple global landmark access condition is likely due to two reasons. First, users only needed one global landmark (the church) for self-localization in the current studies. Second, we realized afterward (from user comments) that users had difficulty in extracting each of the global landmarks from the AR interface, as it was cluttered with too much information, which made it less effective.

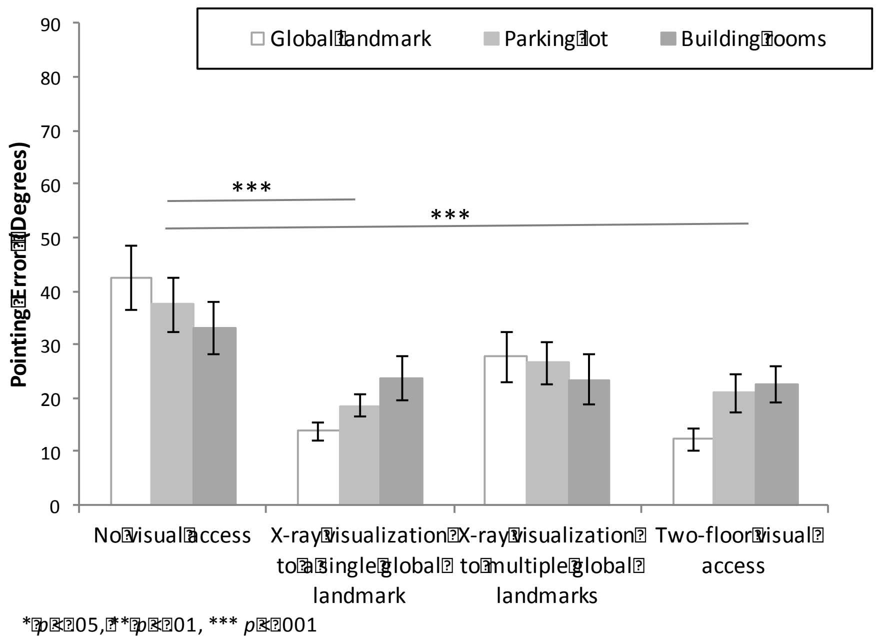


Figure 5.5. Mean absolute pointing error for Experiment 6.

A repeated-measures ANOVA was conducted for each of the three dependent measures of wayfinding accuracy, wayfinding efficiency, and drilling accuracy, with the four conditions of visual access as the within-subject factor. There was no significant main effect of visual access on wayfinding accuracy, $F(3, 189) = .539$, $p > .05$, $\eta^2 = .014$; or wayfinding efficiency, $F(3, 189) = .550$, $p > .05$, $\eta^2 = .009$. These results are consistent with Experiments 4-5. This lack of an effect is likely due to the environments tested, e.g. all buildings in the current studies had congruent floor layouts without any loops and each building consisted of only one elevator. As a result, navigators could find the target room using the shortest path based on accessing two accurate single-floor cognitive maps or even from route knowledge formed during the learning phase. In a previous study, we investigated how the realism of a virtual environment model impacts human wayfinding in a multi-level building (Giudice & Li, 2012). The virtual multi-level building in that study had two elevators and the results showed that a sparsely rendered model significantly promoted users' wayfinding accuracy and efficiency. Thus, we predict that the X-ray visualization used in Experiment 6 could also promote users' wayfinding performance in a complex building with multiple elevators, which will be the topic of a future experiment.

A significant main effect of visual access was observed for the drilling task, $F(3, 189) = 5.548$, $p < .001$, $\eta^2 = .081$. Subsequent Dunn–Sidak pairwise comparisons showed that drilling accuracy in the X-ray visualization (single global landmark access condition) was significantly higher than the no visual access condition ($p < .001$) and the window-access condition ($p < .05$). This is not surprising for the no visual access condition but very meaningful for the window-access condition.

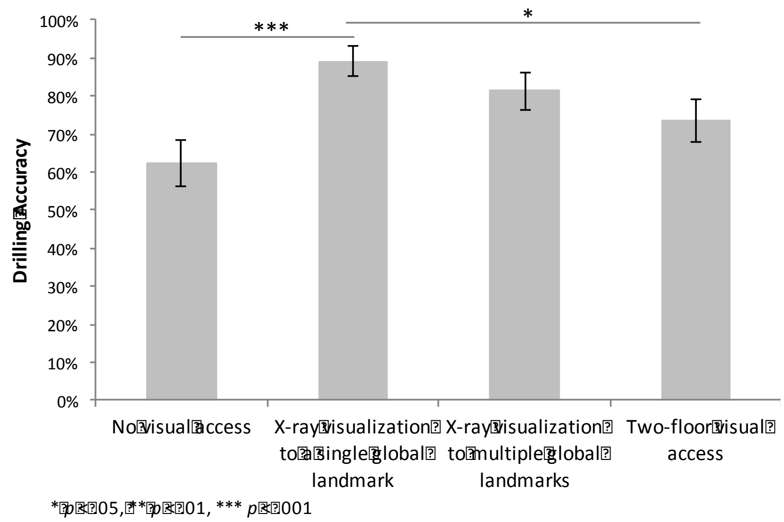


Figure 5.6. Mean drilling accuracy for Experiment 6.

The finding that the X-ray visualization outperformed the window-access condition in promoting users' drilling accuracy suggests that this interface is more than an adequate substitute for the gold standard of windows. Instead, it provided more clear inter-floor visualization than is possible from windows. We interpret this superior pointing and drilling performance as providing evidence that the X-ray visualization affords even better visual access in the multi-level built environment than is possible from observation through the building's windows. Taken together, the results of Experiment 6 provided compelling evidence that the X-ray visualization is an effective approach for promoting users' development of a multi-level cognitive map.

Table 5.2. Mean absolute pointing error, pointing latency, wayfinding accuracy, wayfinding efficiency, and drilling accuracy in Experiment 6.

Dependent Variables	Pointing target type	Visual access			
		No visual access	X-ray visualization to a single global landmark	X-ray visualization to multiple global landmarks	Two-floor visual access
Absolute pointing error	Global landmark	42.51 (6.06)	13.83 (1.65)	27.69 (4.82)	12.26 (2.15)
	Parking lot	37.42 (5.19)	18.52 (2.07)	26.55 (3.91)	20.98 (3.50)
	Building rooms	33.11 (4.98)	23.63 (4.21)	23.45 (4.60)	22.59 (3.51)
Pointing latency	Global landmark	6.56 (0.44)	5.86 (1.08)	6.18 (0.77)	5.07 (0.65)
	Parking lot	8.46 (1.42)	8.42 (1.18)	7.59 (1.05)	6.37 (0.50)
	Building rooms	10.49 (1.23)	11.75 (1.64)	9.37 (1.09)	10.36 (1.20)
Wayfinding accuracy		57.8% (6.2%)	65.6% (6.0%)	59.4% (6.2%)	56.3% (6.3%)
Wayfinding efficiency		55.1% (6.1%)	63.6% (5.9%)	56.2% (5.9%)	56.8% (6.0%)
Drilling accuracy		62.5% (6.1%)	89.1% (3.9%)	81.3% (4.9%)	73.4% (5.6%)

5.3.4 Discussion

On the basis of the findings of Experiments 4-5, we proposed and evaluated a third AR interface in Experiment 6 (an X-ray visualization technique), compared to two control conditions. The most important finding from Experiment 6 is that the X-ray visualization condition outperformed the "gold standard" of window-access in promoting users' development of multi-level cognitive maps, by helping synthesize information across the global landmark and building locations. This finding suggests that increasing visual access with AR techniques is not merely an alternative and economical approach, but a more effective way for overcoming the disadvantage of limited visual access in built environments and improving the development of a multi-level cognitive map. This finding has important practical significance in that the AR technology could make a local landmark that is not physically visible in multiple locations/levels in a building to be a "global" landmark and thereby provide a generalizable, broad-based solution for improving spatial behaviors in complex buildings.

The findings of experiments 5-6 provide three Human-computer interaction principles for designing cognitively motivated visualization techniques for development of indoor navigation systems. First, designers should provide the depth information of the global landmarks on the AR interface by showing transparent walls, occluded hallways, and the horizon. Second, designers should keep the AR visualization uncluttered, i.e.

showing multiple global landmarks is not necessarily helpful. Third, designers should allow users to turn on/off the AR visualization.

5.4 Experiment 7 (2D vs. 3D schematic maps)

5.4.1 Introduction

In Experiment 7, we proposed and evaluated two schematic map visualization approaches (2D map with a top-down viewpoint vs. 3D map with a bird's-eye viewpoint) as shown in Figure 5.7, compared to a control condition (without map assistant). The two visualization approaches are different from the traditional single-floor visualization method in that both show multiple floors at the same time and provide users with access to their alignment (or misalignment) in a straight forward manner.

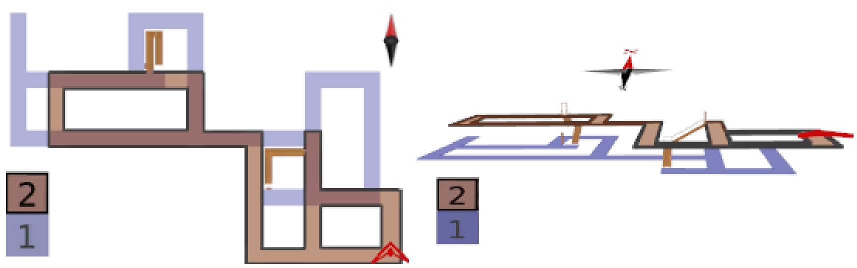


Figure 5.7. A 2D top-down viewpoint map and a 3D bird's-eye viewpoint.

The first research question addressed in Experiment 7 asks whether the map view perspective (2D top-down view map vs. 3D bird's-eye view map) affects the development of multi-level cognitive maps. Several studies have investigated the pros and cons of using 2D and 3D visualizations on portable mobile displays in assisting navigation in both outdoor and indoor spaces (Chittaro & Venkataraman, 2006; Nurminen, 2008; Oulasvirta, Estlander, & Nurminen, 2009). As for outdoor space, Oulasvirta et al. (2009) found users assisted with 2D maps were able to extract more information in less time and used reliable and ubiquitous environmental cues like street names and crossings more frequently than 3D maps; compelling empirical evidence found in their research showed that a 2D street map can outperform a 3D mobile map. However, the 3D map in their research only means a photorealistic representation of the real world (2.5D), as outdoor environments are only composed of a horizontal plane providing no alignment or connectivity information between different planes. In our study, 3D bird's-eye view maps consist of multiple floors/planes and we aim to investigate whether a 3D map could outperform a 2D map for assisting multi-level cognitive map development in these true 3D indoor environments.

Chittaro & Venkataraman (2006) evaluated 2D maps vs. 3D maps for assisting navigation in a multi-level virtual building and the 2D map also outperformed the 3D interface for assisting users to find three object targets whose locations were indicated by the navigation aid. However, in their research, users were assisted by the maps while they were looking for the targets and the target locations were indicated on the maps. In our research, however, participants were only assisted with the maps for visualization in the learning phase, as this study aims to investigate whether the maps facilitated development of multi-level cognitive maps that supported subsequent behavior on cross-floor tasks during testing. Additionally, most users only learned the 3D maps from a fixed viewpoint in the Chittaro & Venkataraman (2006)'s study, as only one participant made use of the sliders to change the map orientation. To solve this problem, in our research, the system automatically rotates the 3D bird's-eye viewpoint when users pass by the targets; this allows users to learn the building from multiple perspectives, as shown in Figure 5.8.

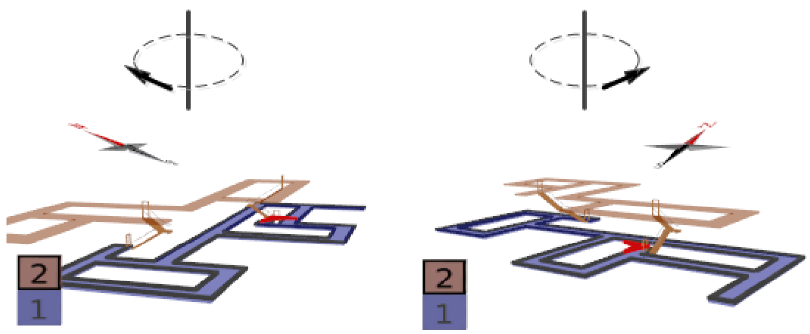


Figure 5.8. Rotation of the 3D bird's-eye viewpoint.

The second research question addressed by Experiment 7 asks how different landmark types (contiguous landmarks that were vertically aligned on each floor and non-contiguous landmarks that had no obvious alignment between floors) might affect performance of cross-level learning and multi-level cognitive map development. Li & Giudice (2012b) proposed the term "contiguous landmark" in the multi-level indoor environment referring to vertically aligned indoor landmarks, which consist of a set of vertically aligned structural or object landmarks located on different floors. For instance, two vertically aligned blue walls can be conceptualized as a contiguous object landmark;

while two vertically aligned T intersections can be conceptualized as a contiguous structural landmark. A constrained field of view of indoor environments causes navigators to have difficulties in learning cross-level visual cues of a building and thereby to depend more heavily on local landmarks, e.g., a picture on the wall, instead of global landmarks (Giudice et al., 2010). Results of Experiments 4-6 showed that access to global landmarks (both from windows-access and from AR visualization) can assist users' ability in forming multi-level cognitive maps. The current study addresses a similar issue but investigates the explicit nature of these local landmarks—whether two vertically-aligned local landmarks could be identified and integrated by the navigator into a global landmark. If this works, it means contiguous landmarks can serve as "glue"—a common spatial reference frame—to align two floors of a building and navigators could use this information as a cue to consolidate local spatial knowledge learned from different floors into a global multi-level cognitive map. Although each part of the contiguous landmark is usually perceived discretely on each floor, the maps make them perceptually available as users can directly visualize the relation between floors and learn the alignment without the normal spatiotemporal constraints and other limitations that usually limit access to cross-floor information integration. Therefore, we propose that contiguous landmarks could help multi-level cognitive map development, as, in principle, they provide a common frame of reference to consolidate the individual floor knowledge into a consistent building-level mental framework.

5.4.2 Methods

Eighteen participants (9 females and 9 males, *M*=25.8, *SD*=7.4) were recruited from the University of Maine.

A Lenovo W510 Thinkpad 15.6-inch workstation notebook with an Intel Core i7 processor and NVIDIA Quadro FX 880M graphics was used. A Logitech Extreme 3D Pro Joystick was used to perform both translational and rotational movements. As shown in Figure 5.9, our environments were comprised of three two-level virtual buildings which were designed using Revit Architecture 2013 (AutoDesk, Inc.). The Unity 4.0 VR engine (Unity Technologies) was used as the VE platform supporting users' real-time navigation and recording their trajectory and test performance.

Figure 5.9. Virtual Environments of Experiment 7.

All three buildings had the same number of corridors and intersections, meaning the layout topology and complexity of all buildings was identical. As shown in Figure

5.10, the solid and dashed line respectively represents the first and the second floor layout. Two vertically aligned T-intersections were used to represent a contiguous structural landmark, while two vertically aligned chandeliers were used to represent contiguous object landmarks. In addition, two non-contiguous single-floor landmarks (an L-intersection and a doorway) were used in each virtual building.

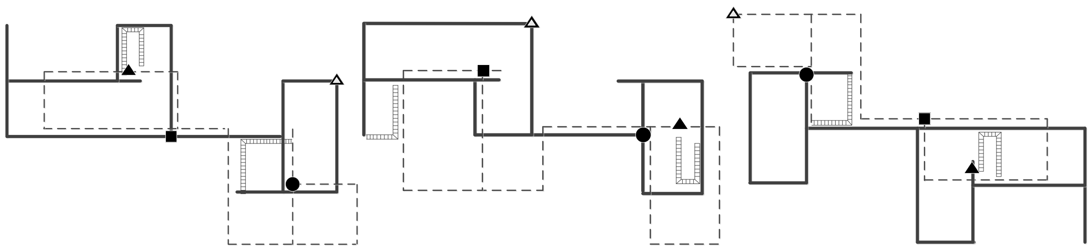


Figure 5.10. Floor layouts of Experiment 7. ● represents target A located at a contiguous object landmark; ■ represents target B located at a contiguous structural landmark; ▲ represents target C located at a doorway; △ represents target D located at an L intersection.

The start learning point was located at the southeast corner of the building. There was a red arrow in the virtual building indicating the start point and the north direction. The two floors were connected by two staircases. There were two pictures on each floor which served as experimental targets. Pictures were based on four high imagery words: chair, table, bottle and clock. The four targets were located at two types of landmarks as shown in Figure 5.10. The targets were initially hidden from view but when participants passed the target, an audio signal was triggered that gave its name. The target also visually appeared for ten seconds and then faded out.

A within-subject design was adopted, with eighteen participants running in three levels of visualization method (2D top-down view map, 3D bird's-eye view map, and a third unaided control condition). Participants first freely learned a multi-level building and then took part in four spatial tasks: pointing, wayfinding, vertical navigation task, and paper-based drilling task (as described in the general experimental procedure in Section 3.3). Finally, participants took part in a user preference survey.

5.4.3 Results

A 3 (visualization method: two map viewpoints and a third no map assistant) × 2 (landmark type: contiguous vs. non-contiguous) repeated-measures ANOVA was conducted for each of the six dependent measures (absolute pointing error, pointing latency, wayfinding accuracy, wayfinding efficiency, vertical navigation accuracy and paper-based drilling accuracy). Significant main effects of visualization method were observed for absolute pointing error, vertical navigation accuracy and paper-based drilling accuracy: absolute pointing error, $F(2, 34) = 42.483$, $p < .001$; vertical navigation accuracy, $F(2, 34) = 3.433$, $p < .01$; paper-based drilling accuracy, $F(2, 34) = 8.859$, $p < .01$. Subsequent Dunn–Sidak pairwise comparisons showed that pointing and vertical navigation in the 2D map and 3D map conditions were more accurate than in the control condition (no map assistance) and paper-based drilling accuracy in the 2D map condition was more accurate than both the 3D map condition and the control condition (all *p*s $< .01$).

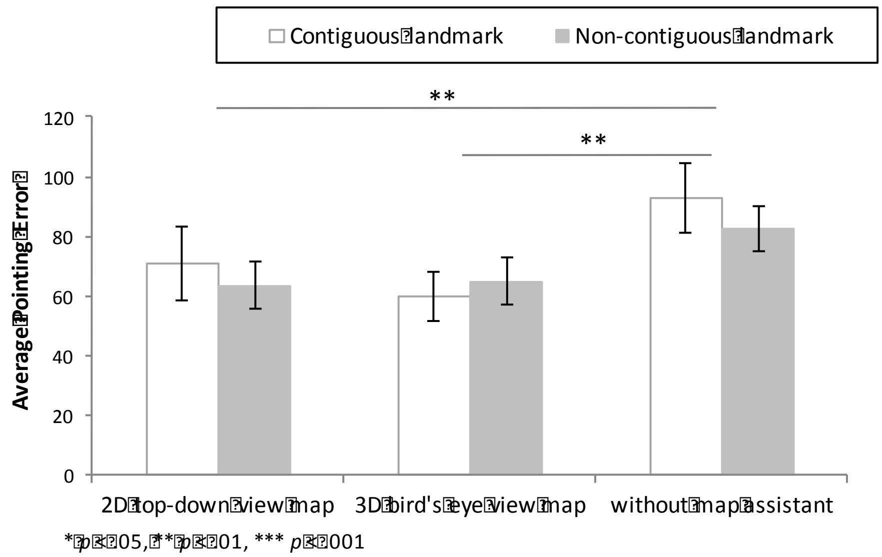


Figure 5.11. Mean absolute pointing error for Experiment 7.

No effects of landmark type were observed for any measure (all *p*s > .10). However, there was a significant interaction between visualization method and landmark type on the paper-based drilling accuracy, $F(2, 34) = 3.923$, $p < .05$. Subsequent Dunn–Sidak pairwise comparisons indicated that users were more accurate at finding targets at contiguous landmarks in the 2D top-down view map condition than in the 3D bird's-eye view map condition ($p < .05$).

Table 5.3. Mean absolute pointing error, pointing latency, vertical wayfinding accuracy, and paper-based drilling accuracy in Experiment 7.

Dependent Variables (Measures)	Landmark type	Visualization method		
		2D map with a top-down viewpoint	3D map with a bird's-eye viewpoint	Without map assistant
Absolute pointing error	Contiguous	70.79 (12.28)	59.89 (8.16)	92.67 (11.68)
	Non-contiguous	63.29 (7.92)	64.92 (7.85)	82.50 (7.35)
Pointing latency	Contiguous	11.49 (1.40)	15.58 (2.68)	14.03 (1.46)
	Non-contiguous	12.70 (1.58)	15.80 (2.45)	11.87 (1.18)
Vertical Wayfinding accuracy	Contiguous	41.7% (7.3%)	44.4% (6.9%)	27.8% (7.3%)
	Non-contiguous	38.9% (8.6%)	41.7% (7.3%)	25.0% (7.3%)
Paper-based Drilling accuracy	Contiguous	30.6% (10.0%)	8.3% (4.5%)	11.1% (5.0%)
	Non-contiguous	19.4% (5.9%)	13.9% (5.4%)	8.3% (4.5%)

5.4.4 Discussion

The primary goal of Experiment 7 was to investigate whether use of either 2D top-down view maps or 3D bird's-eye view maps, compared to no map assistant, significantly improved the development of multi-level cognitive maps. As expected, the reliably better performance on the cross-level spatial tasks with visualization assistance, compared to the control condition, provides evidence of the efficacy of these visualization techniques for promoting the development of multilevel cognitive maps. However, no significant differences were found in the pointing and vertical navigation tasks between the two maps (2D top-down view map vs. 3D bird's-eye view map), which is inconsistent with previous research regarding the evaluation of the 2D and 3D maps (Chittaro & Venkataraman, 2006; Oulasvirta et al., 2009). One possible explanation is that the 3D map used in our research is a true 3D structural rendering rather than a photorealistic representation of the world (Oulasvirta et al., 2009). In addition, users could learn the 3D internal structure of the buildings from multiple perspectives, which is known to be helpful for understanding the internal structure of the object compared to learning from a fixed bird's-eye view (Cohen & Hegarty, 2007), in which, the relative direction and distance between targets can be distorted due to 3D perspective cues. Therefore, the performance gap found between 2D and 3D maps in the previous literature was likely narrowed here, as we made full use of the utility of 3D maps, e.g., by providing multiple perspectives.

Nevertheless, use of 2D maps outperformed 3D maps in the paper-based drilling task, which is an important task to test the development of multi-level cognitive maps (see Sections 2.3 and 3.3). The statistics showed that users had significantly better performance in indicating these targets at contiguous landmarks in the 2D top-down view map condition than the 3D bird's-eye view map condition. This suggests that the 2D maps are a better visualization method for allowing participants to learn the between-floor alignment, which means that they can develop more accurate multi-level cognitive maps supporting behaviors that require cross-floor spatial knowledge. The 2D top-down view map condition consistently outperformed the without map assistant condition in all three tasks. Corroborating the statistical advantage observed for the 2D maps, fourteen participants indicated that they preferred the 2D top-down view map, whereas only four selected the 3D bird's-eye view map as their favorite. No participant chose the without map assistant condition. Some examples of the positive comments include: "the 2D map made the locations easier to see or point out", "2D map is easier to get a sense of the floor layouts in relation to each other. The 3D map was hard to read and learn", "2D map gives me a better idea of vertical navigation and relation of objects". Taken together, these results provide evidence that the 2D top-down view map is a more user-friendly visual interface for assisting multi-level cognitive map development than the 3D map. This preference is not surprising; 2D maps have been used for hundreds, if not thousands, of years. Metro maps, which are very similar to the 2D top-down view maps in this study, are products of decades of research and development.

As for the analysis of contiguous landmarks vs. non-contiguous landmarks analysis, there were no significant differences observed in the pointing and vertical navigation tasks, but there was a significant difference found in the paper-based drilling task. In this study, users received map assistance in the learning phase only and we did not explicitly highlight any contiguous landmarks on either the 2D maps or 3D maps. In addition, users were not informed of any clues about the contiguous landmark information before the experiment. Even with this lack of knowledge, some participants still perceived and made use of these contiguous landmarks for between floor wayfinding, which supports the importance of visualization techniques that highlight and emphasize these cues. Therefore, in the future, we will further investigate whether explicitly highlighting the contiguous landmarks on the maps improve learning of multi-level buildings and subsequent development of multi-level cognitive maps. The findings will provide important guidelines for the design of cognitively motivated visualization techniques for use in the development of indoor navigation systems.

The pointing error and wayfinding accuracy observed in this study are lower than expected even when assisted with the maps for visualization. There are two reasons that we think likely account for this outcome. First, the virtual buildings used in this study have much higher structural complexity for transition between floors than were used in previous research (Richardson et al., 1999). The virtual buildings used in our study have incongruent floor layouts and disorienting staircases, which consist of both horizontal and angular offsets between the floors. Second, the virtual buildings were designed with very low architectural differentiation, which has been discussed as another main cause of getting lost in indoor spaces (Carlson et al., 2010). Even in such a complex indoor environment, participants showed reliably better performance in the two map conditions than the without map assistant control condition, meaning that the proposed two visualization methods are efficient tools in facilitating multi-level cognitive map development.

5.5 Summary

In Chapter 5, we described three experiments (Experiments 5-7) using VEs to evaluate several indoor visualization interfaces, including AR visualizations and schematic maps, for assisting the development of multi-level cognitive maps.

Experiments 5-6 were conducted to evaluate *Hypothesis 4*: we can use AR technology to increase visual access to global landmarks, which could facilitate users' development of multi-level cognitive maps. Hypothesis 4 was validated, as results of Experiment 6 demonstrated that increasing visual access with the X-ray visualization condition is a more effective approach for overcoming the disadvantage of limited visual access in built environments and improving the development of a multi-level cognitive map. Although the results of Experiment 5 showed that the two simply rendered AR models were not effective for facilitating multi-level cognitive map development, the findings provide guidelines for the design of AR-based visualization techniques for use in the development of indoor navigation systems. Experiment 7 was conducted to evaluate *Hypothesis 5*: schematic maps that effectively convey the desired multi-level building information to users could alleviate the challenge of integrating cross-level spatial knowledge. The outcome of Experiment 7 supports Hypothesis 5, since participants had reliably better performance on the cross-level spatial tasks with both maps (2D top-down view map and 3D bird's-eye view map) compared to the no-map condition, suggesting that these visualization techniques could promote the development of multilevel cognitive maps. In sum, the proposed visualization interfaces in this thesis are efficient tools in facilitating multi-level cognitive map development. These findings contribute to the growing field of indoor navigation and provided valuable insights to the real-time indoor navigation systems.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Summary of Contributions

"Space plays a role in all our behavior. We live in it, move through it, explore it, defend it" (O'Keefe & Nadel, 1978). From traveling to school, to driving to work, to flying long distances for vocation, our daily lives involve a myriad of spatial behaviors. This dissertation deals with a special type of space, multi-level buildings, and it addresses a long-standing and ubiquitous problem faced by anyone who navigates indoors—why do people get lost inside multi-level buildings? To investigate this vexing issue, a two-pronged approach was employed combining theories and paradigms from both basic and applied research. This section details the motivation and key findings that came out of my doctoral thesis work.

6.1.1 Contributions Related to Theories of Spatial Cognition

Of theoretical interest, I investigated how multi-level built environments are learned and structured in memory. The most important contribution of this dissertation is the extension of the traditional concept of cognitive maps to multi-level built environments, termed here as multi-level cognitive maps. The concept of multi-level cognitive maps plays an important role in understanding how humans learn and represent multi-level built environments. On the basis of previous behavioural and neurobiological literature regarding human mental representations of 3D space, especially in multi-level built environments, I developed the notion of a multi-level cognitive map, defined as a "multi-layer" structure, consisting of: (1) a set of super-imposed 2D cognitive maps, (2) between-floor connectivity information, (3) between-floor alignment information, and (4) encoding of the between-floor z-axis offset. The concept of a multi-level cognitive map is distinguished from a true 3D mental representation (Yartsev & Ulanovsky, 2013), as the vertical axis of a multi-level cognitive map is not encoded with the same representational structure and fidelity as the horizontal plane. The concept of multi-level cognitive maps is similar to the bicoded three-dimensional spatial encoding model (Jeffery et al., 2013), as both consist of a set of locally planar 2D cognitive maps. However, I studied an important extension of the bicoded theory—how navigators integrate cross-floor spatial knowledge during vertical travel. Compelling empirical evidence was found showing that it is very challenging and error-prone for humans to integrate cross-level spatial knowledge during vertical travel, supporting the theory that people are developing and accessing a bicoded spatial representation rather than a true 3D mental representation. My thesis work contributes important empirical data that advances theories of spatial cognition regarding human spatial learning and mental representations of 3D space.

The findings from nine behavioral experiments validated the framework of multi-level cognitive map development, providing important insights into how humans build a globally coherent mental representation of multi-level buildings.

1) In Experiments 1-3, I successfully disentangled relevant between-floor topological and structural factors that might cause difficulty forming multi-level cognitive maps. Previous studies have suggested that the between-floor z-axis offset is attributable to the challenge people experience in forming a globally coherent mental representation of multi-level buildings, which I posited as being too simplistic an explanation. In support of my claim, results of Experiments 1 and revealed that it is not the presence of the z-axis offset alone but the combination of the between-floor overlap and the z-axis offset that causes the difficulty of integrating cross-level spatial knowledge. Results of Experiment 2 showed that when two floors of a multi-level building are incongruent and misaligned (i.e., the reference directions of the two floors have an angular offset), it is more difficult for people to build a globally coherent mental representation of this building. Results of Experiments 1 and 3 further demonstrated that people perform better in buildings with floors that are aligned and users' ability to form multi-level cognitive maps was not adversely effected or reliably impaired in these environments, even when the orientation of the elevator had an angular offset with respect to the floor's reference direction. These results suggest that if two floors of a

building are aligned and share a common spatial reference frame, navigators can use interior features such as walls or hallways to learn the building structure and grasp the congruent frame of reference between floors. These findings are important for providing new insights into the fundamental spatial cognition question of why it is challenging to integrate cross-level spatial knowledge. More importantly, these findings are specific and concrete enough to provide design guidelines for architects and building planners, which are discussed in the following section.

2) Results of Experiment 4 demonstrated that if a multi-level building affords visual access to a global landmark, navigators can use this global landmark to accurately integrate cross-level spatial knowledge. In addition, if the global landmark is visible from both indoor and outdoor space (O/I-space), it can facilitate the integration of O/I spaces, which is a longstanding challenge discussed in the spatial cognition literature (Giudice et al., 2010). Previous literature has found that global landmarks serving as fixed global spatial references assist navigators when outdoors to integrate local spatial knowledge into a global cognitive map. Findings from Experiment 4 are consistent with this previous research but extend the advantage of accessing global landmarks from single-plane 2D spaces into multi-level built environments. These new results provide important empirical evidence for guiding future spatial cognition research investigating the role of landmarks serving as fixed global spatial references in 3D settings.

3) Despite the clear advantage demonstrated in this dissertation for accessing global landmarks from within buildings, they are often not available in multi-level indoor environments, and it is impractical to modify the physical building to provide this access. However, results of Experiment 5-7 showed that new visualization interfaces can facilitate the integration of cross-level spatial knowledge. For instance, two approaches that were developed and evaluated here and that led to reliably improved spatial performance included augmented reality (AR) visualizations and schematic maps that effectively convey the desired multi-level building information (e.g., global landmarks, floor layouts). Specifically, results of Experiment 6 demonstrated compelling evidence that increasing visual access using an X-ray AR visualization that showed transparent walls, occluded hallways, and the horizon, outperformed the “gold standard” of window-access in promoting users’ development of multi-level cognitive maps. Establishing new and improved visualization techniques are important for spatial cognition research. My thesis research is one of only a few studies investigating the effect of visualization techniques on users’ ability to learn, represent, and behave in multi-level buildings, providing important empirical foundations for future spatial cognition research in this domain.

6.1.2 Contributions Related to Applications and Design

In the previous section, I discussed the findings from Experiments 1-7 contributing to theories in the domain of spatial cognition such as improving our understanding of how a globally coherent mental representation of multi-level buildings is developed to represent and support spatial behaviors in these multi-level environments. In this section, I will describe how the outcomes of my thesis work also significantly contribute to related real-world applications, such as architectural design and the development of cognitively motivated spatial visualizations.

The outcomes of Experiments 1-4 provided four evidence-based design guidelines for architects.

1) Results of Experiments 1 and 2 revealed that it is not the presence of the z-axis offset alone but the combination of the between-floor overlap and the z-axis offset that causes the difficulty of integrating cross-level spatial knowledge.

Guideline 1: based on this knowledge, designers should subdivide each floor of a complex multi-level building into sub-regions, so that sub-regions of two floors are non-overlapped, and thus, the entire building is easier for users to learn and build a multi-level cognitive map. My thesis work only focused on two-floor buildings. However, this logic would hold for more complex multi-level environments (see discussion in Section 6.2.1).

2) Results of Experiment 2 showed that when two floors of a multi-level building are misaligned, it is very challenging and error-prone for people to build accurate multi-level cognitive maps.

Guideline 2: architects should avoid designing buildings with between-floor misalignment, which was found to be a significant cause of people becoming disoriented or lost.

3) Previous studies have suggested that confusing staircases with confusing between floor heading shifts often cause people to get disoriented during indoor navigation. However, results of Experiments 1 and 3 demonstrated that the between-floor heading shift and the portal-floor heading shift did not significantly impair users' ability to form multi-level cognitive maps in aligned buildings. These findings indicate that between-floor alignment is an important factor helping navigators to 'glue' or integrate spatial properties and can even help offset the potential difficulty imposed by confusing between floor heading shifts. In addition, results of Experiment 4 showed that global landmarks can serve as a fixed spatial reference frame for navigators to integrate multi-level spatial knowledge into a multi-level cognitive map.

Guideline 3: if confusing staircases or elevators with misaligned portals are necessary in a building for aesthetic consideration or other reasons, designers should ensure that navigators can easily learn between-floor alignment by providing external cues such as a common between-floor axis (aligned buildings) or situating the staircase/elevator in an open area such as an atrium (providing visual access to global landmarks).

4) Results of Experiment 4 showed that if a global landmark is visible from both indoor and outdoor space, it will facilitate both cross-level spatial knowledge integration and the integration of outdoor and indoor spaces (OI-spaces) into multi-level cognitive maps.

Guideline 4: architects should design buildings with good visual access between indoor and outdoor spaces to help navigators to effectively transit between O/I spaces.

In addition, this thesis work contributes to the growing field of real-time indoor navigation systems and provides new insight into the optimal visualization interfaces to be employed in these systems in order to support the most accurate and intuitive spatial learning and behavior in complex buildings. The findings of Experiments 5-7 provided four HCI principles for cognitively motivated visualization techniques for development of indoor navigation systems.

1) In Experiment 5, I proposed and evaluated two simply rendered AR models (an icon-model and a wireframe-model), which require fewer computational resources to render compared to other visualization techniques. Results showed that the two AR models, although computationally resource-efficient, were not effective for facilitating

multi-level cognitive map development. There is a trade-off between information access and resource efficiency. It makes no sense to design resource efficient visualization techniques that don't actually facilitate the underlying mental representational process. In Experiment 6, I found compelling evidence that increasing visual access using a solution based on an X-ray AR visualization outperformed the "gold standard" of window-access in promoting users' development of multi-level cognitive maps. Taken together, these findings provide three important guidelines that can be used for the design and implementation of real-time AR-based visualization techniques for use in the development of indoor navigation systems.

HCI principle 1: designers should provide information about visual depth of the global landmarks depicted on the AR interface by showing transparent walls, occluded hallways, and the horizon.

HCI principle 2: designers should keep the AR visualization clear and uncluttered (less is more). For example, showing multiple global landmarks on the AR visualization is not necessarily helpful in assisting users to build more accurate multi-level cognitive maps than simply showing one well-chosen global reference.

2) In Experiment 5, users were constantly exposed to the AR information through an always-on interface, which sometimes distracted users from learning the building itself (based on users' comments). Thus, it is important to let users control the on/off of the AR

visualization, e.g. deciding on when they want assistance from presence of the AR visualization.

HCI principle 3, designers should allow users to turn on/off the AR visualization on-demand, rather than adopting an always-on interface.

3) Results of Experiment 7 indicate that both the 2D top-down view map and the 3D bird's-eye view map can promote users' ability to form multilevel cognitive maps. However, the 2D top-down view map was found to be a more effective visualization method for allowing participants to learn the between-floor alignment, which as described earlier, was found in my studies to be a critical factor improving between-floor spatial performance and the development of multi-level cognitive maps.

HCI principle 4: designers should use a 2D top-down view map instead of a 3D bird's-eye view map to best assist users with accurate real-time indoor navigation.

6.2 Future Work

6.2.1 Extending the Concept of Multi-Level Cognitive Maps to More Complex Multi-level Built Environments

The concept of multi-level cognitive maps will be expanded in future studies based on the following four components.

First, in the real world, a complex multi-level building often comprises many aboveground floors and underground levels. If multiple floors of a building have the same (or similar) layouts, these floors might be mentally represented as one level in the multi-level cognitive map. In reality, each floor of a complex multi-level building usually comprises more complex floor layouts than the buildings evaluated in this thesis. Thus, each floor of these buildings might be encoded in multiple regions. The regions of multi-level built environments are likely to be mentally grouped together in psychological space and will thus form superordinate nodes in the multi-level cognitive map. Thus, we postulate that a multi-level cognitive map is not only a “multi-layered” structure but also a regionalized hierarchical structure, meaning that both horizontal and vertical regions could be grouped together and form superordinate nodes in the multi-level cognitive map. The processes involved in mentally representing multiple regions of a complex building in the formation of a multi-level cognitive map will be further investigated in future studies and the findings will be used to enhance the concept of multi-level cognitive maps as advanced in this dissertation.

Second, complex multi-level buildings often contain multiple vertical connectors (e.g., elevators, staircases, escalators). A greater quantity of vertical connectors creates a more complex between-floor connectivity matrix, meaning that it is more challenging to determine the shortest path between positions located on two floors. In addition, using different types of vertical connectors for vertical transitions might affect the development of multi-level cognitive maps. For instance, a horizontal transition shift is often involved in the vertical transition via an escalator but not an elevator in most cases. These issues will be important research topics for future studies, and the findings from this research will be used to refine the concept of multi-level cognitive maps on between-floor connectivity information.

Third, investigating the encoding and distortion of the z-axis (vertical dimension) will be a research topic for a future project. Previous literature has found that humans can roughly estimate distance between floors, but the estimations are distorted with “relative downward errors in upward judgments and relative upward errors in downward judgments” (Jeffery et al., 2013; Tlauka et al., 2007; Wilson et al., 2004). However, in the studies by both Tlauka et al. (2007) and Wilson et al. (2004), each floor of the three-level virtual building contained only one room (vista space). It is necessary to investigate how the encoding of the vertical dimension is distorted in a more complex multi-level building with each floor containing multiple rooms and hallways. These issues will be addressed in future studies, and we will use the findings from these studies to refine the concept of multi-level cognitive maps described in this thesis on the encoding of the z-axis.

Fourth, the work described in this thesis is primarily concerned with “point-to-point” vertical alignment. For instance, in the drilling task, the experimenter asked participants to indicate which object or landmark is directly above/below their current location. Even if the object or landmark is a room, it is perceived and conceptualized as a point or position rather than an area or a region. The reason for use of this experimental design is that the "point-to-point" vertical alignment information is the foundation of the "region-to-region" vertical alignment information. We argue that if users have learned accurate "point-to-point" vertical alignment between floors, they can use this information to infer the "region-to-region" vertical alignment—which region (or regions) are directly above/below their current located region. These issues will be important research topics for future studies to refine the concept of multi-level cognitive maps in regard to between-floor alignment information.

We will use the empirical findings described above to develop a more formal characterization of the concept of multi-level cognitive maps. Hillier and colleagues have developed a methodology of *space syntax* to characterize spatial configuration by describing and analyzing patterns of architectural space at both the building and urban level (Hillier & Hanson, 1984). However, space syntax focuses on the connectivity of single-plane layouts and assigns no special significance to the vertical or 3D qualities of places (Montello, 2007). Evaluating the impact of refined parameters of environmental factors on the development of multi-level cognitive maps may contribute new insights to the space syntax theory by extending it to multi-level indoor spaces. These issues will be addressed in future projects.

6.2.2 Combining AR Visualizations and Schematic Maps to Assist the Development of Multi-level Cognitive Maps

In this thesis, we evaluated two types of visualization interfaces (AR visualizations and schematic maps) for assisting the development of multi-level cognitive maps, as described in Section 2.2.3. Although both types of visualization interfaces were found to facilitate the development of multi-level cognitive maps, the difference between the effects of these visualization interfaces were not directly investigated in this thesis and will be addressed in a future study. In a previous study, we proposed that an important design principle of cognitively motivated visualization techniques for indoor navigation systems is to flexibly combine the advantages of different types of visual interfaces based on users' wayfinding demands. The visual interface we conceptualized was termed the "details-on-demand" visual interface (Li & Giudice, 2012a). The wayfinding demands were categorized according to the user requirements encapsulated in the five navigation phases, as illustrated in Table 6.1.

Table 6.1. Five levels of wayfinding demands for transition between O/I spaces.

Navigation phases	Spaces	Wayfinding demands
Plan to enter a building	O	Overview of the building
Enter a building	O/I	Learn main floor layout and important environmental features such as elevators
Navigation on one floor	I	Learn current floor layout
Navigation between floors	I	Integrating vertical information
Plan to exit a building	I/O	Overview of the spatial references of O/I

The details-on-demand visual interface for facilitating the development of multi-level cognitive maps and the integration between OI-spaces will be further investigated in future research.

REFERENCES

Avelar, S., & Hurni, L. (2006). On the Design of Schematic Transport Maps. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *41*(3), 217–228.

Burt de Perera, T., Holbrook, R., Davis, V., Kacelnik, A., & Guilford, T. (2013). Navigating in a volumetric world: Metric encoding in the vertical axis of space. *Behavioral and Brain Sciences*, *36*(05), 546–547.

Carlson, L. A., Hölscher, C., Shipley, T. F., & Dalton, R. C. (2010). Getting Lost in Buildings. *Current Directions in Psychological Science*, *19*(5), 284–289.

Casakin, H., Barkowsky, T., Klippel, A., & Freksa, C. (2000). Schematic maps as wayfinding aids. In *Spatial Cognition II: Integrating Abstract Theories, Empirical Studies, Formal Methods, and Practical Applications* (Vol. 1849, pp. 54–71). Springer Berlin Heidelberg.

Chittaro, L., & Venkataraman, S. (2006). Navigation aids for multi-floor virtual buildings: A comparative evaluation of two approaches. In *Proceedings of the ACM symposium on Virtual reality software and technology* (pp. 227–235).

Clementini, E., Felice, P. Di, & Hernández, D. (1997). Qualitative Representation of Positional Information. *Artificial Intelligence*, *95*(97), 317–356.

Cohen, C., & Hegarty, M. (2007). Individual differences in use of external visualisations to perform an internal visualisation task. *Applied Cognitive Psychology*, *711*, 701–711.

Dey, A., & Sandor, C. (2014). Lessons learned: Evaluating visualizations for occluded objects in handheld augmented reality. *International Journal of Human Computer Studies*, *72*(10-11), 704–716.

Doeller, C. F., Barry, C., & Burgess, N. C.-P. (2010). Evidence for grid cells in a human memory network. *Nature*, *463*(7281), 657–61 ST – Evidence for grid cells in a human me.

Downs, R., & Stea, D. (1973). Cognitive maps and spatial behavior: Process and products. *Image and Environment*, *8*(26).

Egenhofer, M. J., & Franzosa, R. D. (1991). Point-set topological spatial relations. *International Journal of Geographical Information Systems*, *5*(2), 161–174.

Etienne, A. S., & Jeffery, K. J. (2004). Path integration in mammals. *Hippocampus*, *14*(2), 180–92.

Freksa, C. (1992). Using Orientation Information for Qualitative Spatial Reasoning. *Lncs*, *639*(1), 1–17.

Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., & Moser, M.-B. (2004). Spatial representation in the entorhinal cortex. *Science*, *305*(5688), 1258–1264.

Gallistel, C. R. (1990). *The organization of learning*. The MIT Press.

Garling, T., Böök, A., Lindberg, E., & Arce, C. (1990). Is elevation encoded in cognitive maps? *Journal of Environmental Psychology*, *10*(4), 341–351.

Gärling, T., Lindberg, E., Carreiras, M., & Book, A. (1986). Reference systems in cognitive maps. *Journal of Environmental Psychology*, *6*(1), 1–18.

Giudice, N. A., Betty, M. R., & Loomis, J. M. (2011). Functional equivalence of spatial images from touch and vision: evidence from spatial updating in blind and sighted individuals. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(3), 621–34.

Giudice, N. A., & Li, H. (2012). The effects of visual granularity on indoor spatial learning assisted by mobile 3D information displays. In C. Stachniss, K. Schill, & D. H. Uttal (Eds.), *Spatial Cognition VIII* (Vol. 7463, pp. 163–172). Kloster Seeon, Germany: Springer.

Giudice, N. A., & Tietz, J. D. (2008). Learning with virtual verbal displays: effects of interface fidelity on cognitive map development. *Spatial Cognition VI. Learning, Reasoning, and Talking about Space*, *5248*, 121–137.

Giudice, N. A., Walton, L. A., & Worboys, M. (2010). The informatics of indoor and outdoor space: a research agenda. In *2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness* (pp. 47–53).

Golledge, R. G. (1999). Human wayfinding and cognitive maps. *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*, 5–45.

Goyal, R. K., & Egenhofer, M. J. (1997). The Direction-Relation Matrix : A Representation for Direction Relations between Extended Spatial Objects. *The Annual Assembly and the Summer Retreat of University Consortium for Geographic Information Systems Science*, *3*, 95–102.

Goyal, R. K., & Egenhofer, M. J. (2001). Similarity of cardinal directions. *Advances in Spatial and Temporal Databases*, *2121*(1), 36–55.

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, *436*(7052), 801–6.

Hart, R. A., & Moore, G. T. (1973). The development of spatial cognition: A review. In R. M. Downs & D. Stea (Eds.), *Image and environment* (pp. 246–288). Hawthorne, NY: Aldine.

Hayman, R., Verriotis, M., Jovalekic, A., Fenton, A. A., & Jeffery, K. J. (2011). Anisotropic encoding of three-dimensional space by place cells and grid cells. *Nature Neuroscience*, *14*(9), 1182–1188.

Hegarty, M. (2008). Decoupling of intuitions and performance in the use of complex visual displays. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (Vol. 1, pp. 881–886).

Hegarty, M., Smallman, H. S., Stull, A. T., & Canham, M. S. (2009). Naïve Cartography: How Intuitions about Display Configuration Can Hurt Performance. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *44*(3), 171–186.

Hillier, B., & Hanson, J. (1984). *The social logic of space*. Cambridge University Press.

Hölscher, C., Brösamle, M., & Vrachliotis, G. (2012). Challenges in multilevel wayfinding: A case study with the space syntax technique. *Environment and Planning B: Planning and Design*, *39*(1), 63–82.

Hölscher, C., Büchner, S. J., Meilinger, T., & Strube, G. (2009). Adaptivity of wayfinding strategies in a multi-building ensemble: The effects of spatial structure, task requirements, and metric information. *Journal of Environmental Psychology*, *29*(2), 208–219.

Hölscher, C., Büchner, S., & Strube, G. (2013). Multi-floor buildings and human wayfinding cognition. *Behavioral and Brain Sciences*, *36*(05), 551–552.

Hölscher, C., Meilinger, T., Vrachliotis, G., Brösamle, M., & Knauff, M. (2006). Up the down staircase: Wayfinding strategies in multi-level buildings. *Journal of Environmental Psychology*, *26*(4), 284–299.

Ijsselsteijn, W., de Ridder, H., Freeman, J., Avons, S., & Bouwhuis, D. (2001). Effects of Stereoscopic Presentation, Image Motion, and Screen Size on Subjective and Objective Corroborative Measures of Presence. *Presence: Teleoperators and Virtual Environments*, *10*(3), 298–312.

Jeffery, K. J., Jovalekic, A., Verriotis, M., & Hayman, R. (2013). Navigating in a three-dimensional world. *The Behavioral and Brain Sciences*, *36*, 523–587.

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.

Kant, I., Müller, F. M., & Noiré, L. (1881). *Immanuel kant's critique of pure reason*. London, UK: Macmillan.

Kelly, J. W., & Mcnamara, T. P. (2008). Spatial memories of virtual environments: How egocentric experience, intrinsic structure, and extrinsic structure interact. *Psychonomic Bulletin & Review*, *15*(2), 322–327.

Kim, K., Rosenthal, M. Z., Zielinski, D., & Brady, R. (2012). Comparison of desktop, head mounted display, and six wall fully immersive systems using a stressful task. *2012 IEEE Virtual Reality (VR)*, 143–144.

Kitchin, R., & Blades, M. (2002). *The cognition of geographic space* (Vol. 4). IB Tauris.

Klatzky, R. L., Beall, A., & Loomis, J. M. (1999). Human navigation ability: Tests of the encoding-error model of path integration. *Spatial Cognition and Computation*, *1*, 31–65.

Klatzky, R. L., & Giudice, N. A. (2013). The planar mosaic fails to account for spatially directed action. *Behavioral and Brain Sciences*, *36*(05), 554–555.

Klatzky, R. L., Loomis, J. M., & Golledge, R. G. (1997). Encoding spatial representations through nonvisually guided locomotion: Tests of human path integration. *Psychology of Learning and Motivation*, *37*, 41–84.

Klippel, A., Hirtle, S., & Davies, C. (2010). You-Are-Here Maps: Creating Spatial Awareness through Map-like Representations. *Spatial Cognition & Computation*, *10*(2-3), 83–93.

Klippel, A., Richter, K., Barkowsky, T., & Freksa, C. (2005). The cognitive reality of schematic maps. In L. Meng, A. Zipf, & T. Reichenbacher (Eds.), *Map-based Mobile Services-Theories, Methods and Implementations* (pp. 57–74). Berlin: Springer.

Lau, R. W. H., Li, F., Kunii, T. L., Guo, B., Zhang, B., Magnenat-Thalmann, Nadia Kshirsagar, Sumedha Thalmann, D., & Gutierrez, M. (2003). Emerging web graphics standards and technologies. *Computer Graphics and Applications, IEEE*, *23*(1), 66–75.

Levine, M., Jankovic, I. N., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General*, *111*(2), 157.

Li, H., & Giudice, N. A. (2012a). Details-on-demand Mobile Visual Interface for Facilitating Indoor Wayfinding. In *extended abstract at the conference of GIScience 2012*.

Li, H., & Giudice, N. A. (2012b). Using Mobile 3D Visualization Techniques to Facilitate Multi-level Cognitive Map Development of Complex Indoor Spaces. In C. Graf, N. A. Giudice, & Schmid，F. (Eds.), *Proceedings of the International Workshop on Spatial Knowledge Acquisition with Limited Information Displays (SKALID')* (pp. 31–36). Monastery Seeon, Germany.

Li, H., & Giudice, N. A. (2013). The Effects of 2D and 3D Maps on Learning Virtual Multi-level Indoor Environments. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Map Interaction* (pp. 7–12). Orlando, FL, USA: ACM.

Loomis, J. M., Blascovich, J. J., & Beall, a C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers*, *31*(4), 557–64.

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

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

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*, 125–151.

Lynch, K. (1960). *The Image of the City*. MIT Press.

MacEachren, A. (1992). Application of environmental learning theory to spatial knowledge acquisition from maps. *Annals of the Association of American Geographers*, *82*(2), 245–274.

Mark, D. M., Freksa, C., Hirtle, S. C., Lloyd, R., & Tversky, B. (1999). Cognitive models of geographical space. *International Journal of Geographical Information Science*, *13*(8), 747–774.

McNamara, T. P., Sluzenski, J., & Rump, B. (2008). Human spatial memory and navigation. In H. L. Roediger (Ed.), *Learning and memory: A compre- hensive reference: Vol. 2. Cognitive psychology of memory* (Vol. 2, pp. 157–178). Oxford: Elsevier.

Meilinger, T. (2008). The network of reference frames theory: A synthesis of graphs and cognitive maps. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5248 LNAI, pp. 344–360). Freiburg, Germany.

Meilinger, T., Riecke, B. E., & Bülthoff, H. H. (2013). Local and global reference frames for environmental spaces. *Quarterly Journal of Experimental Psychology*, *67*(3), 542–69.

Montello, D. R. (1993). Scale and multiple psychologies of space. In *Spatial information theory: a theoretical basis for GIS* (pp. 312–321). Springer.

Montello, D. R. (1998). A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In M. J. Egenhofer & R. G. Golledge (Eds.), *Spatial and temporal reasoning in geographic information systems* (pp. 143–154). New York: Oxford University Press.

Montello, D. R. (2007). The contribution of space syntax to a comprehensive theory of environmental psychology. *76th International Space Syntax Symposium Instanbul 2007*, 1–12.

Montello, D. R., & Freundschuh, S. (2005). Cognition of Geographic Information. In McMaster R. B. & E. L. Usery (Eds.), *A research agenda for geographic information science* (pp. 61–91). Boca Raton, FL: CRC Press.

Montello, D. R., & Pick, H. L. (1993). Integrating Knowledge of Vertically Aligned Large-Scale Spaces. *Environment and Behavior*, *25*(3), 457–484.

Moratz, R., & Ragni, M. (2008). Qualitative spatial reasoning about relative point position. *Journal of Visual Languages and Computing*, *19*(1), 75–98.

Moratz, R., & Wallgrün, J. O. (2012). Spatial reasoning with augmented points: Extending cardinal directions with local distances. *Journal of Spatial Information Science*, *5*(5), 1–30.

Moser, E. I., Kropff, E., & Moser, M.-B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annual Review of Neuroscience*, *31*, 69–89.

Mou, W., McNamara, T. P., & Valiquette, C. M. (2004). Allocentric and Egocentric Updating of Spatial Memories. *Cognition*, *30*(1), 142–157.

Nadel, L. (1991). The hippocampus and space revisited. *Hippocampus*, *1*(3), 221–9.

Newton, I. (1687). *Philosophia naturalis principia mathematica*.

Norman, D. A., & Society(U.S.), C. S. (1981). What is cognitive science? In D. A. Norman (Ed.), *Perspectives on cognitive science*. Norwood, N.J. : Ablex Pub. Corp. ; Hillsdale, N.J. : L. Erlbaum Associates.

Nurminen, A. (2008). Mobile 3D City Maps. *IEEE Computer Graphics and Applications*, *28*(4), 20–31.

O'Keefe, J. (1991). An allocentric spatial model for the hippocampal cognitive map. *Hippocampus*, *1*(3), 230–235.

O'Keefe, J., Burgess, N., Donnett, J. G., Jeffery, K. J., & Maguire, E. A. (1998). Place cells, navigational accuracy, and the human hippocampus. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *353*(1373), 1333–40.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171–175.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press.

O'Neill, M. J. (1992). Effects of familiarity and plan complexity on wayfinding in simulated buildings. *Journal of Environmental Psychology*, *12*(4), 319–327.

Oulasvirta, A., Estlander, S., & Nurminen, A. (2009). Embodied interaction with a 3D versus 2D mobile map. *Personal and Ubiquitous Computing*, *13*(4), 303–320.

Papadias, D., & Egenhofer, M. J. (1997). Algorithms for Hierarchical Spatial Reasoning. *GeoInformatica*, *1*, 1–23.

Passini, R. (1992). *Wayfinding in architecture(2nd ed.)*. New York: Van Nostrand Reinhold Company.

Péruch, P., & Gaunet, F. (1998). Virtual environments as a promising tool for investigating human spatial cognition. *Cahiers de Psychologie Cognitive*, *17*(4-5), 881–899.

Polys, N. F., Kim, S., & Bowman, D. A. (2007). Effects of information layout, screen size, and field of view on user performance in information-rich virtual environments. *Computer Animation and Virtual Worlds*, *18*(1), 19–38.

Presson, C. C., DeLange, N., & Hazelrigg, M. D. (1989). Orientation specificity in spatial memory: What makes a path different from a map of the path? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 887–897.

Richardson, A. E., Montello, D. R., & Hegarty, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition*, *27*(4), 741–750.

Richter, K., & Klippel, A. (2005). A model for context-specific route directions. *Spatial Cognition IV. Reasoning, Action, Interaction*, 58–78.

Riecke, B. E., Bodenheimer, B., McNamara, T. P., Williams, B., Peng, P., & Feuereissen, D. (2010). Do we need to walk for effective virtual reality navigation? physical rotations alone may suffice. In C. Hölscher, T. F. Shipley, M. Olivetti Belardinelli, J. A. Bateman, & N. S. Newcombe (Eds.), *Spatial Cognition VII* (pp. 234–247). Springer Berlin Heidelberg.

Ruddle, R. A., Payne, S. J., & 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., Payne, S. J., & Jones, D. M. (1999). Navigating large-scale virtual environments: what differences occur between helmet-mounted and desk-top displays? *Presence*, *8*(2), 157–168.

Sadalla, E. K., & Montello, D. R. (1989). Remembering Changes in Direction. *Environment and Behavior*, *21*(3), 346–363.

Schmid, F. (2008). Knowledge-based wayfinding maps for small display cartography. *Journal of Location Based Services*, *2*(1), 57–83.

Schmid, F., Richter, K.-F., & Peters, D. (2010). Route Aware Maps: Multigranular Wayfinding Assistance. *Spatial Cognition & Computation*, *10*(2-3), 184–206.

Schultheis, H., & Barkowsky, T. (2013). Just the tip of the iceberg: The bicoded map is but one instantiation of scalable spatial representation structures. *Behavioral and Brain Sciences*, *36*(05), 565–566.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in Child Development and Behavior*, *10*, 9–55.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *... : Teleoperators and Virtual Environments*, 1–20.

Smallman, H. S., Cook, M., & Cowen, M. (2007). Naïve realism in terrain appreciation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1–6.

Smallman, H. S., & John, M. S. (2005). Naive Realism: Misplaced Faith in Realistic Displays. *Ergonomics in Design: The Quarterly of Human Factors Applications*, *13*(3), 6–13.

Soeda, M., Kushiyama, N., & Ohno, R. (1997). Wayfinding in Cases with Vertical

Motion. In *Proceedings of MERA 97: Internatinal Conference on Environment-Behavior Studies* (pp. 559–564).

Stanton, D., Wilson, P., & Foreman, N. (1996). Using virtual reality environments to aid spatial awareness in disabled children. In *Proceedings of the 1st European conference on disability, virtual reality and associated technologies* (pp. 93–101).

Steck, S. D., & Mallot, H. a. (2000). The Role of Global and Local Landmarks in Virtual Environment Navigation. *Presence: Teleoperators and Virtual Environments*, *9*(1), 69–83.

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology*, *10*(4), 422–37.

Street, W. N. (2012). *Between Floor Navigation Cost Depends on Visual Information*. University of Illinois at Urbana-Champaign.

Tlauka, M., & Wilson, P. N. (1996). Orientation-free representations from navigation through a computer-simulated environment. *Environment and Behavior*, *28*(5), 647–664.

Tlauka, M., Wilson, P. N., Adams, M., Souter, C., & Young, A. H. (2007). An Investigation into Vertical Bias Effects. *Spatial Cognition & Computation*, *7*(4), 365–391.

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

Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology*, *13*(3), 407–433.

Tversky, B. (1992). Distortions in cognitive maps. *Geoforum*, *23*(2), 131–138.

Ulanovsky, N. (2011). Neuroscience: How is three-dimensional space encoded in the brain? *Current Biology*, *21*(21), R886–R888.

Uzzell, D. (1995). The myth of the indoor city. *Journal of Environmental Psychology*, 299–310.

Vidal, M., Amorim, M.-A., & Berthoz, A. (2004). Navigating in a virtual three-dimensional maze: how do egocentric and allocentric reference frames interact? *Brain Research. Cognitive Brain Research*, *19*(3), 244–58.

Vidal, M., & Berthoz, A. (2005). Navigating in a Virtual 3D Maze: Body and Gravity, Two Possible Reference Frames for Perceiving and Memorizing. *Spatial Cognition & Computation*, *5*(2), 139–161.

Waller, D., Montello, D. R., Richardson, A. E., & Hegarty, M. (2002). Orientation specificity and spatial updating of memories for layouts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1051–1063.

Wan, X., Wang, R. F., & Crowell, J. a. (2013). Effects of Basic Path Properties on Human Path Integration. *Spatial Cognition & Computation*, *13*(March 2015), 79–101.

Wang, R. F., & Street, W. N. (2013). What counts as the evidence for three-dimensional and four-dimensional spatial representations? *Behavioral and Brain Sciences*, *36*(05), 567–568.

Ware, J. M., Taylor, G. E., Anand, S., & Thomas, N. (2006). Automated Production of Schematic Maps for Mobile Applications. *Transactions in GIS*, *10*(1), 25–42.

Werner, S., & Schindler, L. E. (2004). The Role of Spatial Reference Frames in Architecture: Misalignment Impairs Way-Finding Performance. *Environment & Behavior*, *36*(4), 461–482.

Wiener, J. M., & Mallot, H. (2003). "Fine-to-Coarse"Route Planning and Navigation in Regionalized Environments. *Spatial Cognition and Computation*, *3*(4), 331–358.

Wiener, J. M., Schnee, a, & Mallot, H. A. (2005). Use and interaction of navigation strategies in regionalized environments. *Journal of Environmental Psychology*, *24*, 475–493.

Wilson, P. N., Foreman, N., Stanton, D., & Duffy, H. (2004). Memory for targets in a multilevel simulated environment: evidence for vertical asymmetry in spatial memory. *Memory & Cognition*, *32*(2), 283–97.

Wright, T., & Madey, G. (2008). A survey of collaborative virtual environment technologies. *University of Notre Dame-USA, Tech. Rep*, 1–16.

Yang, L. (2015). *Theories and Models of Indoor Space*. the University of Maine.

Yartsev, M. M., & Ulanovsky, N. (2013). Representation of Three-Dimensional Space in the Hippocampus of Flying Bats. *Science*, *340*(6130), 367–372.

APPENDIX

GLOSSARY

Alignment effect: when the reference direction of a map (visual/haptic) is misaligned with respect to the navigators' facing direction in the environment, judging the direction of environmental features represented on the map is slower and less accurate than when the map is aligned with their facing direction.

Allocentric reference system: the location and orientation of objects are specified with respect to the environment.

Bearing*:* the direction from the navigator to a landmark, measured with respect to the reference direction.

Between-floor heading shift: the angular offset between a navigator's facing directions at a pair of transition points (denoted by α). It is also called the transition angular offset.

Between-floor misalignment: If two floors of a building have a perspective shift γ, they can be said to have a *misalignment* between the floors.

Between-floor overlap: if two floors of a building are overlapped, there must be a set of positions within the two floors co-located at the same x-y coordinates.

Between-floor portal: portal of a vertical connector (e.g., an elevator door or a stairway), through which navigators enter or exit a floor via a vertical connector.

Bicoded representational structure: the spatial representation in surface–traveling animals, comprising a mosaic of these locally planar bicoded map fragments rather than a fully integrated volumetric map, such that space in the plane of locomotion is represented differently from space in the orthogonal axis.

Cognitive map: an enduring, observer-free spatial representation of the environment.

Course*:* the direction of a navigator's velocity vector, measured with respect to the reference direction.

Decision point: the point where two route-segments meet or the intersection of two or more corridors / travel paths.

Egocentric reference system: the objects and spatial relationships of the environment are organized with respect to the observer's position and orientation.

Environmental space: building-sized spaces perceived by moving through the space.

Figural space: object-sized spaces perceived from one vantage point.

Geographical space: large scale outdoor spaces experienced from symbolic representations, such as maps.

Global landmarks: distant landmarks such as towers or mountain peaks that are visible from a large field of view / area in the environment.

Global reference system: spatial reference frame for the whole environmental space.

Heading: the direction of a navigator's facing direction, measured with respect to a reference direction.

Horizontal transition offset: the offset between the transition point and the projection of the corresponding transition point on the former transition point's floor.

Landmarks: distinctive objects or scenes stored in memory.

Local landmarks: landmarks visible only from a limited distance or perceptual field of view.

Local reference system: spatial reference frame for local spaces.

Multi-level cognitive maps: a globally coherent mental representation of multi-level built environments. This term is largely interchangeable with another term multi-level survey knowledge in this thesis. A multi-level cognitive map is constructed by integrating single-level spatial knowledge learned from different floors (local reference systems) into a common spatial reference frame (global reference system). It consists of (1) a set of super-imposed 2D cognitive maps, (2) between-floor connectivity information (e.g., elevators, staircases, escalators, etc.), (3) between-floor alignment information (e.g., indicating what is directly above/below one's current location), and (4) encoding the z-axis offset (e.g., rough estimates of floor heights).

Multi-level landmarks: distinctive objects or scenes that are visible from multiple locations/levels of a multi-level built environment. In this thesis, multi-level landmarks are interchangeable with global landmarks.

Multi-level route knowledge: the knowledge of travel paths that connect between-floor locations. Multi-level route knowledge is based on an egocentric spatial reference frame.

Multi-level survey knowledge: a configurational representation of the metric spatial relationship between environmental features across multiple floors organized within a common spatial reference frame. See multi-level cognitive maps.

Path integration: the process of obtaining rotational displacements by updating heading information during vertical transition.

Perspective shift: the angular offset that is necessary to move from one reference frame to the next. It consists of both a translation and a rotation component. The rotation component is denoted as γ.

Piloting: the use of configurations of landmarks to determine one's location or heading.

Portal-floor heading shift: angular offset between the reference direction of a floor and a navigator's heading when entering/exiting a between-floor portal, denoted as β.

Reference direction: the orientation of a reference frame.

Region: perceived and encoded representations in spatial memory in which locations are grouped within a common spatial reference frame and form super-ordinate nodes.

Route: a trace in the environment of a traveled sequence of path segments and turn angles that are followed in order to get from an origin to a destination.

Route knowledge: the knowledge of travel paths that connect landmarks.

Spatial reference frames: a relational system that consists of reference objects, located objects, and the spatial relations that may exist among them. This is also called a spatial frame of reference.

Spatial reference systems: a relational system that consists of reference objects, located objects, and the spatial relations that may exist among them.

Survey knowledge: a configurational representation of spatial relationships between non-linearly-aligned sets of environmental features such as routes and landmarks, organized within a common spatial reference frame.

Transition angular offset: the angular offset between navigators' facing directions at a pair of transition points. It is also called the between-floor heading shift (denoted by α).

Transition point: a point where navigators enter or exit a floor.

Vertical connectors: elevators, staircases, escalators, etc., supporting users in navigating between floors/levels.

Vertical transition: navigators use vertical connectors (e.g., elevators, staircases, escalators, etc.) to navigate between floors.

Vertical transition offset: the z-axis offset between a pair of transition points located at different floors.

Vista space: room-sized spaces perceived from one vantage but allowing for head rotation.

Z-axis offset: vertical distance between floors. A multi-level building contains multiple floors and each floor has a z-axis value (e.g. floor height), meaning that different floors have a z-axis offset.

BIOGRAPHY OF THE AUTHOR

Hengshan Li was born in Suizhou, Hubei Province in China on December 7, 1981. Hengshan received his B.S. in Information Engineering from Wuhan University (Wuhan, China) in 2003, and his M.S. in Cartography and Geographic Information Systems (GIS) from Wuhan University (Wuhan, China) in 2006. From 2006 to 2010, Hengshan started to work as a GIS software engineer and then became a project leader in the Geographic Information Center of Zhejiang (Hangzhou, China). He won the Excellent Programmer of the Geographic Information Center of Zhejiang in 2007 and the 2nd place GIS project award of GIS Association of China in 2008.

In August 2010, Hengshan began the Ph.D. program in the Department of Spatial Information Science and Engineering, University of Maine, under the supervision of Dr. Nicholas Giudice. He was a graduate research assistant funded by the National Science Foundation project (IIS-0916219, entitled: Information integration and human interaction for indoor and outdoor spaces) (2010-2013). He received the University of Maine Chase Distinguished Research Assistantship (CDRA) (2013-2014) and the University of Maine Janet Waldron Doctoral Research Fellowship (UMDRF) (2014-2016). Hengshan also received the Third Place Oral Presentation Award in the University of Maine Graduate Academic Expo (2015) and the University of Maine 2015 Fall Graduate Student Government (GSG) grant.

During his Ph.D. study, Hengshan has published in *Spatial Cognition*, *International Conference on Human-Computer Interaction, Cartographica: The International Journal for Geographic Information and Geovisualization, ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness, International Workshop on Spatial Knowledge Acquisition with Limited Information Displays,* and *ACM SIGSPATIAL International Workshop on Map Interaction.* Hengshan is a reviewer of *Cartographica* and *Future Internet*.

Hengshan Li is a candidate for the Doctor of Philosophy degree in Spatial Information Science and Engineering from the University of Maine in May 2016.