The effects of immersion and body-based rotation on learning multi-level indoor virtual environments

Conference Paper · November 2013

DOI: 10.1145/2533810.2533811

2 authors:

Hengshan Li
Future Cities Laboratory (FCL)

Nicholas A Giudice
University of Maine

The Effects of Immersion and Body-based Rotation on Learning Multi-level Indoor Virtual Environments

Hengshan Li
School of Computing and Information Science
University of Maine
Orono, Maine 04469, USA
hengshan.li@umit.maine.edu

Nicholas A. Giudice
School of Computing and Information Science
University of Maine
Orono, Maine 04469, USA
nicholas.giudice@maine.edu

ABSTRACT

The goal of this study was to investigate how the immersion level of virtual environments (HMD vs. desktop) and rotation method (physical vs. imagined) affects wayfinding performance in multi-story virtual buildings and the development of multi-level cognitive maps. Twelve participants learned multi-level virtual buildings using three VE conditions (physical rotation HMD, physical rotation desktop and imagined rotation desktop). They were then tested on four cross-level tasks, including: pointing, route navigation, vertical navigation, and paper-based drilling. Results showed that performance on between-floor trials was reliably worse than for within-floor trials and that this difference was neither improved by the level of immersion of the display nor the rotation behavior used during navigation. Our data suggest that increasing the fidelity of these interface variables does not yield more accurate development of multi-level cognitive maps. Indeed, multi-level indoor wayfinding performance was as effective with the simplest and least expensive desktop display based purely on joystick navigation as the more complex VE platforms. These findings show that spatial cognition research in multi-level virtual buildings need not be limited to immersive VEs with physical body rotation which require considerable equipment cost and increased technical complexity.

Categories and Subject Descriptors

H.1.2 [**User/Machine Systems**]: Human factors; H.5.1 [**Multimedia Information Systems**]: Artificial, augmented, and virtual realities; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces

General Terms

Design, Experimentation, Human Factors

Keywords

Indoor wayfinding, multi-level cognitive map, immersive virtual environments, body-based rotation

ISA 13, November 05-08 2013, Orlando, FL, USA

http://dx.doi.org/10.1145/2533810.2533811

1. INTRODUCTION

Virtual environments (VEs) have been used in many domains relating to indoor spaces, including for emergency response training scenarios [2], by architects to provide virtual walk-throughs [1, 4, 29], and by spatial cognition researchers studying spatial learning and wayfinding within buildings (see [15] for review). The advantage of virtual environment technology (VET) is that it readily facilitates manipulation of building layout and information content, as well as tracking of navigators' movement behavior. In addition, previous studies have shown that people can ultimately develop accurate spatial knowledge from exposure to large-scale VEs and that this knowledge is similar to knowledge acquisition gained from physical environments [18, 22, 27, 28], although the learning process in VEs typically takes a longer time than in physical buildings [20, 22]. Owing to the ubiquity of portable navigation devices, there is growing interest in researching user experiences, behavioral strategies, and spatial learning using digital indoor maps to navigate large buildings [9]. However, there are still many questions about the efficacy of this technology for learning and navigating multi-level (i.e. two or more story) buildings, whether virtual or physical.

An important characteristic of most complex indoor spaces, e.g. airports, libraries, and office buildings, is that they consist of multiple floors, which often cause navigators to become disoriented or lost when traversing between levels. For instance, Holscher et al. identified incongruent floor layouts, disorienting staircases, and lack of visual access to important level-related building features as the main causes of these wayfinding difficulties when traveling in complex multi-level buildings [8]. Several studies have shown that navigators are significantly less accurate when pointing to locations between floors than within a single floor and poor inter-floor knowledge has been argued as a major cause of disorientation in multi-level buildings [6, 8, 26]. Given the aforementioned literature highlighting the difficulties of inter-level wayfinding behavior, there is a surprising dearth of research into the underlying theory of why integrating multi-level building information is so challenging for human spatial cognition. Therefore, further investigation of spatial behavior in complex multi-level indoor spaces is an important step in forming a theoretical framework of indoor wayfinding, which is at the heart of the current paper. Furthermore, a better understanding of the contributions and underlying mechanisms for how multi-level cognitive maps (the spatial representation of the multi-level building) are developed will benefit the future design of interactive virtual maps used in indoor navigation systems to support the most accurate spatial behaviors. Although the majority of research on multi-level indoor wayfinding has been conducted in physical environments [7, 8, 10, 17], a small body of research has also been conducted in multi-level indoor VEs [3, 6], findings that motivate the current experiments.

2. Research Questions

The first research question addressed by this paper asks how the immersion level of the virtual environment affects wayfinding performance and 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" [25]. This research question is addressed by comparing highly immersive VEs where the information is presented through a head-mounted display (HMD) and changes with head movement vs. low immersion desktop VEs, where the information is presented on a monitor. Compare 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 these VEs and obtain a sense of presence that is not possible using desktop VEs (see [15] for review). Therefore, there is a common assumption that they are more effective than desktop VEs for certain spatial behaviors (e.g., searching, spatial learning, 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 [30] and are therefore rarely used outside of the research lab setting. In addition, previous literature has shown that the HMD-based systems often elicit higher levels of simulator sickness and evoke larger negative emotions compared to desktop VEs [11]. By contrast, desktop VEs have become less expensive and the requisite graphics cards have become more powerful, meaning that the purchase and use of high quality desktop VE environments is more prevalent than in the past [12]. Hence, if performance with desktop VEs is shown to be similar to that with an HMD in some circumstances, use of this simpler technology may open the door for their use in a much broader application domain by many more people, instead of being limited to a few research labs as is generally the case with immersive VE systems. It is therefore critical to assess whether there is sufficient benefit to justify using the more costly immersive VE systems in spatial cognition research. Studies have shown that immersive VEs are more effective than desktop VEs in some close-range tasks, e.g. target detection ([19], see [15, 24] for review). However, other studies have provided evidence that desktop VEs are as effective as HMD-based VEs in some tasks. For instance, Mizell, et al. [16] found no significant differences between the HMD and desktop VE when they were used to learn a three-dimensional sculpture and were subsequently asked to make a physical model of it. As for relevant studies of navigation in VEs, Ruddle et al. [24] 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 participants using the HMD navigated the buildings significantly more quickly. However, 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 [24]. However, the VEs used in the aforementioned experiments were not based on multi-level buildings, which are important characteristics of indoor space. As we are interested in elucidating the theoretical underpinnings of multi-level cognitive maps, it is important that we render multi-level buildings. Furthermore, none of the previously discussed studies have investigated how use of learning in virtual environments (immersive or desktop) affects participants' cross-level spatial knowledge, e.g., the absolute locations and relative directions of targets between floors and the corresponding point above or below a known landmark on the current floor. To perform such tasks with the greatest efficiency, we postulate that users must access a multi-level cognitive map developed by integrating knowledge of each floor when learning the building. In this paper, we propose that the development of multi-level cognitive maps plays an important role for supporting accurate performance on a range of tasks that require integrating knowledge between floors, as well as the more commonly studied within-floor measures. What is unknown from the literature is how the immersion level of the virtual environment affects wayfinding performance and the development of multi-level cognitive maps.

The second research question addressed in this paper asks whether the rotation method used in the desktop VE (imagined rotation desktop vs. physical rotation desktop) affects wayfinding performance and the development of multi-level cognitive maps. Navigation is the most common interactive task performed with wayfinding experiments in VEs. Unfortunately, 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. [21] conducted a study in which participants were asked to search through a computer generated environment for targets in a joystick condition, walking condition (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 and Tietz [5] conducted a study to investigate the effect of physical body rotation using virtual verbal displays (without HMDs) for environmental learning and wayfinding; the results also showed that employing physical rotation during learning significantly improved spatial knowledge acquisition and cognitive map development. 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 rotational component of movement in desktop VEs, especially a multi-level desktop VE, is still unknown. To our knowledge, there has been no research comparing the effectiveness of physical rotation vs. imagined joystick rotation on wayfinding using desktop VEs, as we did in this research. If the results show that imagined rotation in desktop VEs is effective in supporting multi-level wayfinding, the setup of these VE systems will be further simplified, as only a joystick will be necessary and use of inertial tracking to update users' heading during physical rotation will be eliminated.

In a nut shell, this paper focuses on investigating users' performance in learning and navigating multi-level virtual buildings and employs a host of tasks requiring cross-level building knowledge to assess the development of multi-level cognitive maps. Three within-subject conditions (physical rotation HMD, physical rotation desktop and imagined rotation desktop) were used. To address the question of immersion, analysis will be done between the physical rotation HMD and physical rotation desktop conditions, as the only difference between these conditions is the level of VE immersion. To address the role of rotation, analysis will be done between the two desktop conditions, as they are matched on immersion level (desktop VE) but differ only on the rotation parameter.

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. METHOD

3.1 Participants

Twelve participants (6 female and 6 male, mean age=25.9, SD=6.7) were recruited from the University of Maine student body. 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

In the immersive HMD VE condition, we used a zSight integrated SXGA HMD (Sensics, Inc), incorporating inertial tracking, 70 degree 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 (information content was matched between displays). A Logitech Extreme 3D Pro Joystick was used in all three conditions to perform forward-back translations. In the imagined rotation desktop condition, participants use the joystick to make both translational and rotational movements. In all conditions, users sat on a rotatable chair with an attached platform (68*50 cm) 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 1. 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.

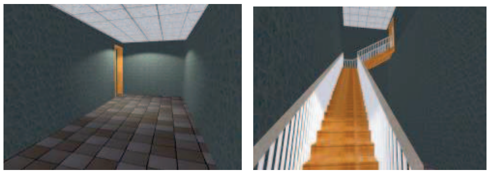


Figure 1. Virtual Environments (multi-level building).

As shown in Figure 2, the solid line represents the first floor layout and the dashed line represents the second floor layout. All virtual buildings used in the experiment were matched for layout complexity and topology, with both floors based on the same bounding rectangle (20*40 meters) with one a dead-end branch. As previous research has indicated that incongruent floor layout is a cause of difficulty in indoor wayfinding [8], the current VEs were designed with 74% overlap of the two floors. Users could move continuously in the virtual building by the joystick at a fix speed (6m/s forward/backward, 3m/s sideways). For analytic convenience, the floor was broken into 5 m segments.

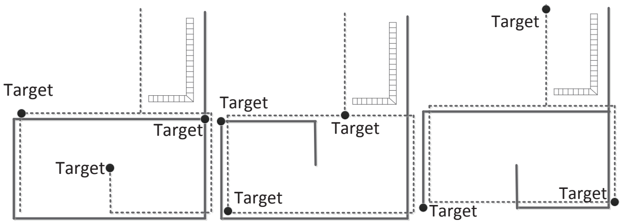


Figure 2. Experimental layouts with target locations denoted.

The start learning point was located at the southeast corner of the first floor. There was a red arrow in the virtual building 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. An additional practice building was used with another three pictures: table, bottle and dog. 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 appeared visually for ten seconds and then faded out.

3.3 Procedure

A within-subject design was adopted, with twelve participants running in all three conditions. There were six phases in the experiment.

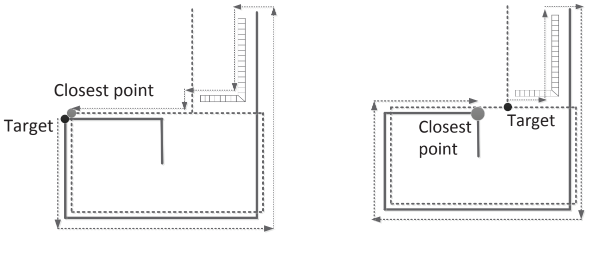
Phase 1: *Practice*. Participants were familiarized with the apparatus and navigation behavior in the VE. All experimental tasks were explained and demonstrated before starting the experimental trials.

Phase 2: *Environmental learning*. From a north orientation at the learning start point, participants freely learn the building for 3.5 minutes (established as sufficient from pilot studies) by means of a user-defined open search. They were asked to find all three targets and remember their relative direction to each other and the building's configuration as a whole. If users found all three targets early in their exploration, they were encouraged to keep on learning the environment until their time had expired. All participants successfully found all three targets within the limited learning period. The only dependent variable analyzed during the learning phase was the number of targets hit, based on the sum of times participants navigated to each individual target.

Phase 3: *Pointing task*. Participants were set to one of the target locations in the virtual building and asked to turn to face another target. They were instructed to face the direction that made a straight-line between the starting and ending target pair, ignoring the walls or paths to walk the route between them. To perform the task, participants must rotate in place and pull the joystick's trigger when they thought they were facing the correct direction to the destination target. If the targets were located on different floors, they were asked to ignore the height offset and point as if the targets were located on the same plane as the floor they were currently on. This task tested whether participants successfully learned the three target locations from phase 2 and could situate them in a globally coherent cognitive map of the building. Accurate pointing required them to make Euclidean judgments from one target to the other target. There are two between-floor pointing trials and one within-floor pointing trial. The two dependent variables analyzed were pointing time and pointing error.

Phase 4: *Navigation task*. Participants were automatically set at one of the target locations and asked to navigate to another target location using the shortest path. Both time and error were measured and participants were asked to navigate as quickly as possible without sacrificing accuracy. In this phase, the target is hidden, so when participants reached the target, they must pull the joystick's trigger to indicate that they had reached the requested location from memory. The order of the navigation routes was counterbalanced. Three dependent variables were analyzed for the navigation task. The first was navigation accuracy, based on whether participants successfully navigated to the correct location of the target. The second was navigation efficiency, based on whether the shortest route was executed (e.g., shortest route length over traveled route length). If incorrect, a third measure of navigation error was calculated, based on the route distance (segments) between the subject's response location and the location of the physical target.

Phase 5: *Vertical Navigation task*. Participants were set to 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 they were currently located at floor 1 at target chair, the task would be to navigate to the corresponding point on floor 2 that is directly above chair, as shown in Figure 3 (a). In each building, there was one target that was located at a place where there is no corresponding point on the other floor's corridor. In this case, they were asked to navigate to the location that was closest to the corresponding vertically-aligned point, as shown in Figure 3 (b).

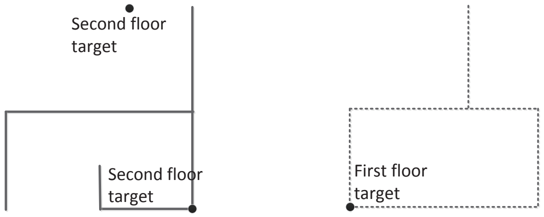


(a)Target and Closest point are vertically aligned (b)Target and Closest point are not vertically aligned

Figure 3. Routes (represented by the dashed lines with arrows) between the target and the closest point to the corresponding vertically-aligned point.

Two dependent variables were analyzed for the vertical navigation task. Vertical navigation accuracy was based on whether participants successfully navigated to the corresponding point of the target. Vertical navigation error distance was based on the Euclidean distance (segments) between subject's response location and the physical location of the corresponding vertically-aligned target.

Phase 6: Paper-based drilling task. In this task, participants were first given a paper which showed the first floor layout. The experimenter provided the targets' names of the second floor. The task was to draw circles on the first floor layout to indicate the imagined vertical locations of second-floor targets on the depicted first-floor map, as shown in Figure 4 (a). Next, participants were given the second floor layout and asked to indicate the first floor targets using the same procedure as shown in Figure 4 (b). As the vertically-aligned targets may or may not correspond to the corridor structure of the depicted map, participants were told that targets could be outside of the map.



(a)First floor layout (b)Second floor layout

Figure 4. Drawing the targets located on the other floor.

Two dependent variables were analyzed for the drilling task. The first was drilling accuracy, based on whether subject successfully drew the targets' corresponding point. The second was drilling distance, based on the Euclidean distance (segments) between subject's drilling location and the actual location of the corresponding vertically-aligned target.

4. RESULTS

4.1 Learning phase

The number of targets hit during exploration were subjected to a one-way repeated measures ANOVA having three levels of VE condition (physical rotation HMD, physical rotation desktop and imagined rotation desktop). The main effect of VE condition was significant, $F(2, 70) = 4.436$, $p < .015$. To address the immersion level question, a pairwise comparison of the physical rotation HMD condition with the physical rotation desktop condition was performed, revealing a significant outcome, $t(35) = -2.547$, $p<.015$, indicating that the average number of targets hit was significantly less in the higher immersion condition (HMD) (M=2.1, SE=.15) than in the lower immersion condition (desktop) (M=2.7, SE=.20). Similarly, to address the role of rotation, a pairwise comparison between the two desktop conditions was performed but was not significant, $t(35) = -.572$, $p >.571$.

4.2 Pointing Task

The pointing time was subjected to a two-way repeated measures ANOVA having three levels of VE conditions and two levels of floor (within-between floor trials). The main effect of floor was significant, $F(2, 22) = 8.209$, $p < .015$, indicating that the average pointing time was reliably longer for the between-floor targets (M=17.3, SE=3.0) than for the within-floor targets (M=9.6, SE=2.1). The main effect of VE condition was non-significant, $F(2, 22) = 1.144$, $p > .337$ and there were no interaction effects between floor and VE condition.

Another two-way repeated measures ANOVA on pointing error was also run for the two factors. The main effect of floor was significant, *F*(2, 22) = 36.496, *p* < .0001, indicating that the average pointing error was reliably higher for the between-floor targets (*M*=53.0, *SE*=6.3) than for the within-floor targets (*M*=15.4, *SE*=3.3). There was no significant effect of VE condition, *F*(2, 22) = 1.274, *p* > .299 or any interaction effects.

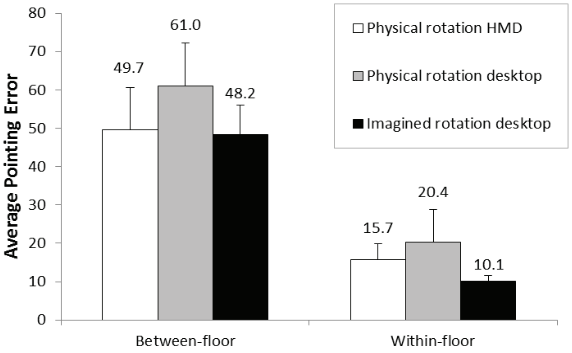


Figure 5. Comparison of average pointing error (± SEM) between three levels of VE conditions and two levels of floor.

4.3 Navigation Task

A two-way repeated measures ANOVA on navigation accuracy was run for the three VE conditions and floor factor. The main effect of floor was significant, *F*(2, 22) = 5.077, *p* < .046, indicating that the average navigation accuracy was significantly lower for the between-floor targets (*M*=69.4%, *SE*=7.9%) than for the within-floor targets (*M*=86.1%, *SE*=5.0%). The main effect of VE condition was not significant, *F*(2, 22) = 1.900, *p* > .173 and there were no reliable interaction effects.

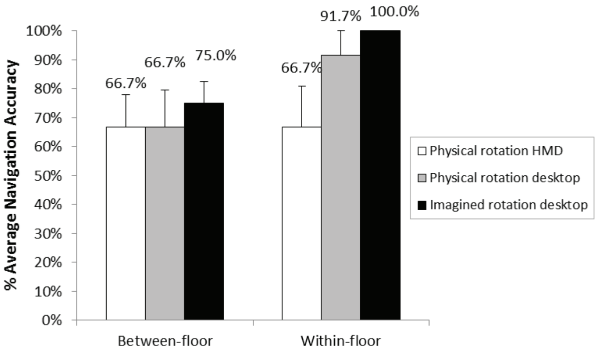


Figure 6. Comparison of average navigation accuracy (± SEM) between three levels of VE conditions and two levels of floor.

A two-way repeated measures ANOVA on navigation efficiency was run for the same factors. Neither the main effect of floor or VE condition was significant, *F*(2, 22) = 3.203, *p* > .101 and *F*(2, 22) = 1.036, *p* > .372 respectively.

A two-way repeated measures ANOVA on navigation error was run for the same factors. The main effect of floor was significant, *F*(2, 22) = 7.383, *p* < .02, indicating that the average navigation error was significantly higher for between-floor targets (*M*=1.6 segments, *SE*=.6) than for within-floor targets (*M*=.8 segments, *SE*=.4). The main effect of VE condition was not significant, *F*(2, 22) = 1.877, *p* > .177.

Taken together, these findings indicate that performance on the navigation task was almost identical between all three VE conditions, and that as predicted, the between-floor navigation trials were reliably worse than the analogous within-floor trials across all conditions.

4.4 Vertical Navigation Task

A one-way repeated measures ANOVA on vertical navigation accuracy and navigation distance error were run for the three VE conditions. Neither the main effect of vertical navigation accuracy or navigation distance error was significant, *F*(2, 70) = .836, *p* > .438 and *F*(2, 70) = .055, *p* > .946 respectively.

In this vertical navigation task, users navigated from one target to the corresponding vertically-aligned point, while in the previous navigation task, users navigated routes between targets. To investigate whether there was a significant effect of the task type on users' navigation accuracy, a one-way repeated measures ANOVA on navigation accuracy was run for the two tasks (route navigation task vs. vertical navigation task). The main effect of task type was significant, *F*(1, 107) = 27.913, *p* < .0001, indicating that users yielded significantly higher target localization accuracy in the route navigation task (*M*=75.0%, *SE*=4.2%) than in the vertical navigation task (*M*=41.7%, *SE*=4.8%), showing that participants could more accurately navigate routes between targets on different floors than aligning a target with its corresponding vertically-aligned point on a different floor.

4.5 Paper-based Drilling Task

A one-way repeated measures ANOVA on drilling accuracy was run with the three levels of VE condition. The main effect was significant, *F*(2, 70) = 3.237, *p* < .045. Addressing the immersion level question, a pairwise post hoc comparison of the physical rotation HMD condition with the physical rotation desktop condition was significant, *t*(35)=-2.256, *p*<.030, indicating that the average drilling accuracy was significantly worse in the higher immersion condition (HMD) (*M*=25.0%, *SE*=7.3%) than in the lower immersion condition (physical rotation desktop) (*M*=47.2%, *SE*=8.4%). To address the role of rotation, a pairwise post hoc comparison of the two desktop conditions was conducted and also revealed significant differences, *t*(35)=-2.092, *p* <.044, indicating that the average drilling accuracy in the physical rotation desktop condition (*M*=47.2%, *SE*=8.4%) was significantly higher than in the imagined rotation desktop condition (*M*=25.0%, *SE*=7.3%).

A one-way repeated measures ANOVA on drilling distance error was run with the same factor. The main effect of VE condition was not significant, *F*(2, 70) =.961, *p* > .388.

5. DISCUSSION

The primary motivation of this study was to investigate whether desktop VEs are as effective as more immersive HMD-based VEs for supporting learning of multi-level virtual buildings and subsequent development of multi-level cognitive maps, as evaluated by a host of behavioral test measures comparing within-floor vs. between-floor performance. The current results provide evidence that the desktop VE is similarly efficient for supporting users' navigation in multi-level virtual buildings. No significant differences were found in the pointing, navigation, and vertical navigation tasks between the two VE immersion levels (HMD vs. desktop). Indeed, in the pointing and navigation tasks, the performance in the desktop VE condition was numerically better than in the HMD VE condition.

There may still be advantages to use of immersive VEs, but the results of this paper indicate that at least with our tasks, the benefit gained from increasing immersion (e.g., HMD) may not be as pervasive as has been suggested in the literature. The most likely reason is that in this experiment, users had to navigate upstairs or downstairs in order to form a multi-level cognitive map of the building. The HMD VE platform cannot provide a vivid illusion of vertical travel in the stairway and hence, it is likely that higher immersion did not provide a benefit over the desktop VE and therefore led to no performance advantage. Second, participants in the immersive condition needed to wear an HMD on their head for about 15 minutes (including the learning and testing phases) which has been shown to sometimes cause discomfort for participants [11]. In corroboration, several participants in this experiment self-reported dizziness caused by wearing the HMD. As a result, the potential advantages of higher immersion may unfortunately be offset by limitations of the equipment on such factors. Where our results do not necessarily indicate that the desktop VE with a joystick is definitively the best interface for learning and navigating multi-level environments, they do suggest that this interface is an effective alternative for supporting these behaviors until cheaper, lighter and more comfortable immersive VEs are developed.

As for the second factor of rotation method (physical vs. imagined) in desktop VEs, the results showed that there were no significant differences 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 here than for the other two conditions, which limits definitive claims about the advantage of physical rotation from these data. The results found in this study are different from previous VE literature [5, 21, 23] in that the inclusion of physical rotation did not lead to a general trend for significantly improved performance. This finding is likely related to several factors. First, in this study, the role of rotation was addressed by analyzing two desktop VE conditions (physical rotation desktop vs. imagined rotation desktop). To execute rotations, users sat on a rotatable chair and turned in place. It is possible that this method of rotation was less intuitive and natural than rotating in place while standing as was used in previous studies that found advantages for physical rotation in virtual environments [5, 21]. In addition, when users rotated in the desktop VE conditions, they still had to fixate directly ahead on the laptop's screen; therefore, there were less degrees of freedom for rotation than in the immersive VEs, where head movement was unrestricted. Therefore, the potential advantages of physical rotation from proprioceptive and vestibular feedback may have been reduced in the current design. Second, the "spiral" movement along the narrow staircases requires good navigation skills in the VE. Several participants in the physical rotation conditions had difficulty in navigating smoothly along the staircases, as they usually made a turn too early or too late and thereby became stuck by the stair railing. In the physical rotation conditions, to execute smooth turns along the staircases, users had to simultaneously turn in the real world and move forward using the joystick. Thanks to the Logitech Extreme 3D Pro Joystick used in the imagined rotation desktop condition, users only needed to use one device to execute both rotational and translational movements 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 presence better than standard joysticks or keyboards used in previous literature [21, 23]. The vertical travel, providing important translation and rotation information [14], is unfortunately the hardest aspect for virtual navigation in the physical rotation desktop VE, so it is perhaps not surprising that the performance on the test measures in the imagined rotation desktop condition (with joystick) was no worse and sometimes numerically better than in the physical rotation desktop VE condition.

In summary, this study did not provide any reliable and consistent evidence to suggest that more immersive VE systems and use of physical body rotation leads to better performance. However, the results showed that participants exhibited greater errors, and had lower navigation accuracy across all conditions when pointing and navigating to between-floor targets than within-floor targets. This result is consistent with previous literature studying multilevel indoor navigation [6, 8, 26]. To emphasize, although there were floor effects, there were no reliable interaction effects in any of the tasks we tested. This means that people were worse for between-floor trials irrespective of the VE condition they used. Therefore, we propose that neither increased immersion nor increased rotational fidelity helps users construct more accurate multi-level cognitive maps. Some situations may benefit from higher immersion or physical movement but at least for the range of tests used in the current complex multi-level environments, the desktop VR with joystick is sufficient and does not warrant the use of more complex equipment, especially given that HMD VEs are less comfortable, more expensive, harder to implement, and have increased technical complexity.

Finally, the results showed that the vertical navigation performance from one target to the corresponding vertically aligned point is harder than navigation between targets, as for all VE conditions, users found significantly more correct targets in the route navigation task (75.0% localization accuracy) than in the vertical navigation task (41.7% localization accuracy). Although users need to travel to another floor in the navigation task, it is essentially a route task as the spatial knowledge required was reinforced in the learning phase. The vertical navigation task, however, is testing the multi-level cognitive map, by evaluating whether participants can vertically align positions on different floors that were never explicitly specified as routes. These corresponding points' locations were not landmarks, as they were in the target-to-target navigation task. It is possible that during training participants traveled the route from one target to its vertically-aligned position but there was no reason to code this in memory. Hence, the routes between the target and the vertically aligned corresponding point have to be aligned based on accessing an accurately formed multi-level cognitive map. Our results indicate that it is very challenging to create such a multi-level cognitive map and to integrate vertical knowledge into a coherent 3D representation of the building. Future studies need to look at other methods to improve the development of multi-level cognitive maps beyond considering the VE technology or interface fidelity. We propose that the relatively poor performance on between floor trials observed here was due to poor visual access to inter-floor information. Preliminary support for this hypothesis has been obtained, showing that better visualization of the layered structure of the building (e.g., as from dynamic 2D overlapped maps) lead to more accurate multi-level cognitive map development [13].

6. ACKNOWLEDGMENTS

This research was supported by NSF grant number: IIS-0916219.

7. REFERENCES

[1] Anderson, L. et al. 2003. A virtual environment for conceptual design in architecture. *Proceedings of the workshop on Virtual environments*. (2003), 57–63.

[2] Chen, Y.-F. et al. 2008. The use of virtual world platforms for supporting an emergency response training exercise. *CGAMES conference*. 3, (2008).

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

[4] Drettakis, G. et al. 2007. Design and evaluation of a real-world virtual environment for architecture and urban planning. *Presence: Teleoperators and Virtual Environments*. 16, 3 (2007), 318–332.

[5] Giudice, N.A. and 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, (2008), 121–137.

[6] Giudice, N. A. and Li, H. 2012. The Effects of Visual Granularity on Indoor Spatial Learning Assisted by Mobile 3D Information Displays. *Spatial Cognition VIII*. C. Stachniss et al., eds. Springer Berlin Heidelberg. 163–172.

[7] Hölscher, C. et al. 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 (2009), 208–219.

[8] Hölscher, C. et al. 2006. Up the down staircase: Wayfinding strategies in multi-level buildings. *Journal of Environmental Psychology*. 26, 4 (Dec. 2006), 284–299.

[9] Hölscher, C. et al. 2013. Wayfinding Cognition & Mobile Maps for Indoor Settings. *www-users.cs.umn.edu*. (2013).

[10] Hölscher, C. and Brösamle, M. 2007. Capturing indoor wayfinding strategies and differences in spatial knowledge with space syntax. *6th International Space Syntax Symposium* (Istanbul, 2007).

[11] Kim, K. et al. 2012. Comparison of desktop, head mounted display, and six wall fully immersive systems using a stressful task. *2012 IEEE Virtual Reality (VR)*. (Mar. 2012), 143–144.

[12] Lau, R.W.H. et al. 2003. Emerging web graphics standards and technologies. *Computer Graphics and Applications, IEEE*. 23, 1 (2003), 66–75.

[13] Li, H. and Giudice, N.A. 2013. The Effects of 2D and 3D Maps on Learning Virtual Multi-level Indoor Environments. *1st ACM SIGSPATIAL International Workshop on Interacting with Maps (MapInteract 2013)* (Orlando, USA, 2013).

[14] Li, H. and Giudice, N.A. 2012. Using Mobile 3D Visualization Techniques to Facilitate Multi-level Cognitive Map Development of Complex Indoor Spaces. *SKALID 2012: Workshop on Spatial Knowledge Acquisition with Limited Information Displays* (Bavaria, Germany., 2012).

[15] Loomis, J.M. et al. 1999. Immersive virtual environment technology as a basic research tool in psychology. *Behavior research methods, instruments, & computers*. 31, 4 (Nov. 1999), 557–64.

[16] Mizell, D.W. et al. 2002. Comparing immersive virtual reality with other display modes for visualizing complex 3D geometry. *Boeing and University College* (London, 2002).

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

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

[19] Pausch, R. et al. 1993. A user study comparing head-mounted and stationary displays. *Virtual Reality, 1993. Proceedings., IEEE 1993 Symposium on Research Frontiers in* (1993), 41–45.

[20] Richardson, A.E. et al. 1999. Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition*. 27, 4 (Jul. 1999), 741–750.

[21] Riecke, B.E. et al. 2010. Do we need to walk for effective virtual reality navigation? physical rotations alone may suffice. *Spatial Cognition VII*. C. Hölscher et al., eds. Springer Berlin Heidelberg. 234–247.

[22] Ruddle, R. a. et al. 1997. Navigating buildings in "desk-top" virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied*. 3, 2 (1997), 143–159.

[23] Ruddle, R. and Lessels, S. 2006. For efficient navigational search, humans require full physical movement, but not a rich visual scene. *Psychological Science*. 17, (2006), 460–465.

[24] Ruddle, R.A. et al. 1999. Navigating large-scale virtual environments: What differences occur between helmet-mounted and desk-top displays? *Presence : Teleoperators and Virtual Environments*. 8, 2 (1999), 157–168.

[25] Slater, M. and Wilbur, S. 1997. A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and virtual environments*. 6, 6 (1997), 603–616.

[26] Soeda, M. et al. 1997. Wayfinding in Cases with Vertical Motion. *In Proceedings of MERA 97: Intern. Conference on Environment-Behavior Studies* (1997), 559–564.

[27] Stanton, D. et al. 1996. Using virtual reality environments to aid spatial awareness in disabled children. *Proceedings of the 1st European conference on disability, virtual reality and associated technologies* (1996), 93–101.

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

[29] Ulicny, B. and Thalmann, D. 2001. Crowd simulation for interactive virtual environments and VR training systems. *Computer Animation and Simulation 2001*. (2001), 163–170.

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