

# Assessment of Visualization Interfaces for Assisting the Development of Multi-level Cognitive Maps

Hengshan Li, Richard R. Corey, Uro Giudice, and Nicholas A. Giudice<sup>(✉)</sup>

Spatial Informatics Program: School of Computing and Information Science, University of Maine,  
348 Boardman Hall, Orono, ME 04469, USA  
Hengshan.li@umit.maine.edu,  
{richard.corey,nicholas.giudice}@maine.edu, uro@vemilab.org

**Abstract.** People often become disoriented and frustrated when navigating complex, multi-level buildings. We argue that the principle reason underlying these challenges is insufficient access to the requisite information needed for developing an accurate mental representation, called a multi-level cognitive map. We postulate that increasing access to global landmarks (i.e., those visible from multiple locations/floors of a building) will aid spatial integration between floors and the development of these representations. This prediction was investigated in three experiments, using either direct perception or Augmented Reality (AR) visualizations. Results of Experiment 1 demonstrated that increasing visual access to a global landmark promoted multi-level cognitive map development, supporting our hypothesis. Experiment 2 revealed no reliable performance benefits of using two minimalist (icon-based and wire-frame) visualization techniques. Experiment 3, using a third X-ray visualization, showed reliably better performance for not only a no-visualization control but also the gold standard of direct window access. These results demonstrate that improving information access through principled visualizations benefit multi-level cognitive map development.

**Keywords:** Multi-level indoor wayfinding · Multi-level cognitive maps · Human factors · Visualization interface design · X-ray visualization

## 1 Introduction

Most travelers can recall an unpleasant memory of becoming disoriented when navigating inside a large building. These buildings usually have a complex multi-story structure with many levels and confusing staircases. Getting lost wastes our time and energy, not to mention being stressful and frustrating. It is widely accepted that to efficiently find our destination in complex environments without becoming lost, navigators rely on the support of cognitive maps—an enduring, observer-free spatial representation of the environment [1, 2]. 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, which has been termed a multi-level cognitive map [3, 4]. Multi-level cognitive maps are postulated as consisting of: (1) a set of super-imposed single-level 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 of the z-axis (e.g., rough estimates of floor heights) [3]. The notion of multi-level cognitive maps of complex built environments is different from the concept of a true 3D spatial representation (see reviews in [5–7]), as the vertical axis of a multi-level cognitive map is not encoded with the same representational structure and fidelity as the x, y axis [3]. Although previous literature has found evidence that the hippocampus can represent 3D volumetric space using a uniform and nearly isotropic rate code along three axes, as with Egyptian fruit bats [6, 8], no evidence for such 3D representations has been observed in humans. By contrast, Jeffery and colleagues [7] suggested a bico-coded representational structure—where 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 bico-coded map fragments rather than a fully integrated volumetric map” [7]. Indeed, there has been a lively debate concerning the efficacy of this bico-coded representation. However, little hard evidence is available to support whether humans were born with the capacity to construct true 3D spatial representations in the brain [7, 9–11]. The consensus is that humans have the capability to encode elevation and z-axis offset in both outdoor and indoor spaces, even if not in a precise 3D manner [12, 13]. For instance, previous studies have found clear evidence that differences in terrain elevation are encoded in cognitive maps of outdoor environments [12]. With regard to indoor environments, a growing body of evidence also suggests that the integration of multi-level spatial knowledge (learned from different floors) can be consolidated into a multi-level cognitive map, but this process is challenging and error-prone for humans to perform [4, 13–16]. Addressing this challenge, the primary goal of the current work is to investigate whether increasing visual access to a global landmark from within a multi-level building could facilitate users' development of a multi-level cognitive map.

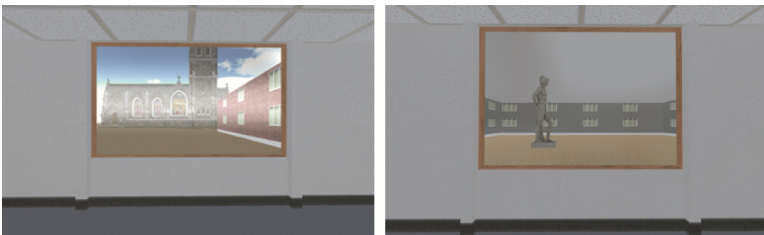
Global landmarks are salient environmental features visible at a large spatial scale from within the environment, e.g., a prominent building. Previous literature on outdoor wayfinding has found clear evidence that these global landmarks provide a fixed spatial reference frame for navigators to integrate local spatial knowledge into a global cognitive map (see [17] for review). However, there is no empirical evidence on the effect of global landmarks observed from within a building in supporting the development of multi-level cognitive maps. This issue is evaluated in Experiment 1. In the three experiments discussed here, users' development of multi-level cognitive maps are measured by three cross-level spatial tasks including pointing and wayfinding between floors and a cross-floor drilling task. The present research also aims to investigate whether visual access to global landmarks can facilitate users' integration of outdoor and indoor spaces, called OI-spaces, which has attracted increasing attention in recent years (see [18] 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.

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 reference frame—is often greatly reduced when learning and navigating through indoor environments [18]. 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 without direct visual access to these 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 [19]. If we can use AR technology to increase visual access to global landmarks, as is investigated in Experiments 2 and 3, 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. All experiments discussed in this article were conducted using virtual environments (VEs), as VEs best facilitate manipulation of building layout and information content, as well as tracking of movement behavior.

## 2 Experiment 1

We propose that a global landmark, serving as a fixed global spatial reference, helps users to consolidate single-level spatial knowledge into a consistent/global multi-level cognitive map. Thus, our hypothesis in Experiment 1 is that users would develop a more accurate multi-level cognitive map when they could see the global landmark from both floors of the experimental building rather than from only a single floor. As shown in Fig. 1, 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.



**Fig. 1.** Outdoor and indoor global landmarks

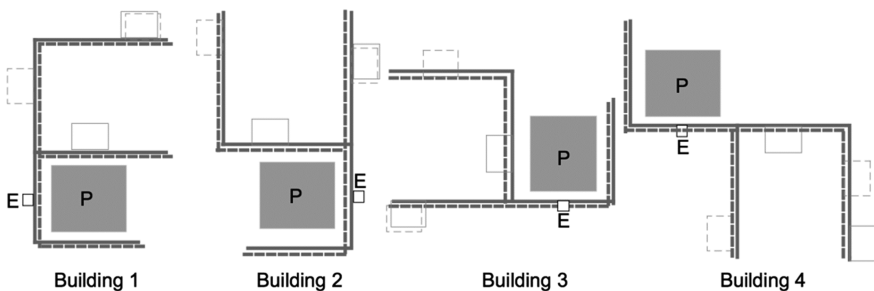
In a previous study, we investigated whether two vertically-aligned chandeliers co-located on separate floors, called contiguous landmarks, could serve as a global landmark and facilitate users’ development of a multi-level cognitive map [4]. However, we observed no reliable effects of contiguous indoor landmarks and very few users even

noticed that the chandeliers were vertically aligned. We interpreted this 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, we predict that both a statue in an atrium and an external landmark, as shown in Fig. 1, can serve as a global landmark, as they are directly perceivable from multiple locations/levels in the building. This assertion was evaluated in the current study.

## 2.1 Method

**Participants:** Sixteen participants (eight females and eight males, mean age = 20.1, SD = 2.0) 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.

**Materials and Apparatus:** The experimental environments were displayed on a Samsung 43" Class Plasma HDTV monitor running at 60 Hz and at a resolution of  $1024 \times 768$ . The desktop VEs were run with a MacBook Pro (2.2 GHz Intel Core i7). The Unity 5.1 VR engine (Unity Technologies, <http://unity3d.com>) 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 Fig. 2. Participants used an elevator to move between floors. All buildings were matched for layout complexity and topology.



**Fig. 2.** Floor layouts. The (*solid line*) represents the first-floor layout. The (*dashed line*) represents the second-floor layout. “E” represents the elevator. “P” represents the parking lot.

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 located throughout the building, as shown in Fig. 3. 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 Fig. 3, each floor consisted of 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 Fig. 3. 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.

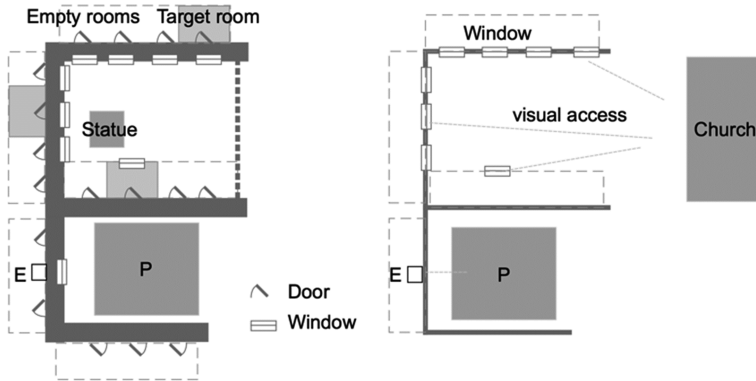


Fig. 3. Visual access to the indoor and outdoor global landmark

## 2.2 Procedure

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). There were five phases in the experiment.

Phase 1: Practice. Subjects 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: Learning. At the beginning of the experiment, participants were positioned at the parking lot. A red arrow on the ground indicated north. Participants were asked to turn in-place and to note the presence/location of the global landmark (e.g., church or statue) from this position. Participants were then guided by blue arrows on the ground to learn the whole building. When they passed by a target room, an audio signal was played that indicated its name, e.g. conference room.

Phase 3: Pointing criterion task. This task was designed to test whether participants had successfully learned the four target rooms. They were first randomly positioned at the doorway of one room and a red arrow appeared to indicate north. The experimenter then asked them to look around and use what they could see of the building's layout,

along with the provided north arrow, to get oriented. When participants were ready, they walked to the center of the room and turned to face north, indicated by the red arrow. The experimenter then 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 degrees. If they failed the first iteration, the Phase 2 learning and Phase 3 pointing criterion tests proceeded until they either successfully met criterion or until they made five incorrect attempts. All participants passed the criterion test within five iterations ( $M = 1.5$ ).

Phase 4: Pointing task. Participants were first randomly positioned at the doorway of a room and were told its name, e.g., "you are facing the conference room". They were encouraged to orient themselves as they did in Phase 3. The experimenter then gave a target room name and asked them to turn to face a straight line to that target. If the target room was on a different floor, they were instructed to ignore the height offset 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 oriented so as to indicate a straight line to the requested target. The experimenter then asked them to point to the global landmark and the parking lot in the same manner. 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 the doorway of a room and received self-orientation as they did in Phase 3. They were then given one target room name and asked to navigate to it using the shortest possible route. Upon reaching the door where they believed the target was located, they turned to face it and pulled the joystick's trigger. The door opened 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: navigation accuracy (whether participants indicated the correct location and orientation of the target room) and navigation efficiency (shortest route length over traveled route length).

Phase 6: Drilling task. After participants had entered a room in the wayfinding task (above), 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). The drilling task tested whether participants successfully learned between-floor alignment information, which is an important component of the multi-level cognitive map.

## 2.3 Results and Discussion

The five dependent measures (pointing latency, absolute pointing error, navigation accuracy, navigation efficiency, and drilling accuracy) were analyzed for each participant. A 2 (visual access: single-floor vs. two-floor)  $\times$  2 (global landmark type: indoor vs. outdoor)  $\times$  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,  $F(1, 63) = 10.057$ ,  $p = .002$ ,  $\eta^2 = .138$ . 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 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  $ps < .001$ ). 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 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  $ps < .05$ ).

A 2 (visual access)  $\times$  2 (global landmark type) repeated-measures ANOVA was conducted for each of the three dependent measures of navigation accuracy, navigation 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,  $F(1, 63) = 4.817$ ,  $p = .032$ ,  $\eta^2 = .071$ . There were no significant main effects of visual access (all  $ps > .172$ ) or global landmark type (all  $ps > .242$ ) on navigation accuracy or navigation efficiency.

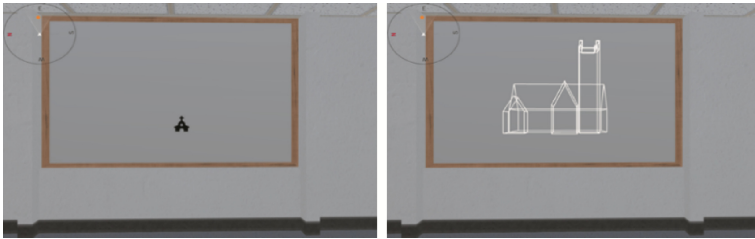
In Experiment 1, we investigated whether increasing visual access to an indoor or outdoor global landmark observed through the building's windows would assist users' development of a multi-level cognitive map. As we predicted, the results demonstrated that users' pointing was reliably faster and more accurate in the two-floor visual access condition than in the single-floor visual access condition, providing clear evidence that a global landmark (both indoor and outdoor) can serve as a fixed spatial reference frame for navigators to integrate multi-level spatial knowledge into a globally coherent multi-level cognitive map. These findings provide important empirical foundations for the design of Augmented Reality (AR) models used in Experiments 2 and 3, 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.

With respect to the variable of global landmark type (indoor vs. outdoor), the results showed that the indoor global landmark was as efficient as the outdoor global landmark for promoting users' pointing between building rooms and pointing to the global landmark. However, results also demonstrated that the outdoor global landmark yielded better pointing performance than the indoor global landmark when pointing to the parking lot, suggesting that an outdoor reference is better in facilitating users' integration between indoor and outdoor spaces. This finding is likely due to the nature of indoor global landmarks, which are often not visible from the outdoor space (e.g., the statue in the atrium was not readily visible from the parking lot in the current study). By contrast, an outdoor global landmark is often visible from both

indoor and outdoor spaces. We believe that the difference found in integrating these environments is due to this disparity in information access and would be eliminated if the indoor and outdoor global landmarks had the same visual access from both within and outside the building. This prediction will be evaluated in a future project. There was a small 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 effect of visual access on drilling accuracy was not observed, meaning that two-floor visual access to a global landmark was not more efficient than single-floor access for promoting drilling accuracy. Indeed, we believe that drilling accuracy may have been elevated for all conditions in Experiment 1 because the fire extinguishers and water fountains were always located directly above/below a target room and some participants indicated that they used this as a cue. This issue is addressed in Experiment 2.

### 3 Experiment 2

The results of Experiment 1 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, Experiment 2 aimed to use AR technology to extend the benefits found in Experiment 1 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 Fig. 4.



**Fig. 4.** Icon-model and wireframe-model 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 Fig. 4. 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. The two AR models require fewer computational resources to render and take less time to create when compared to other visualization techniques, as reviewed in [19]. 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.

Sixteen new students participated in Experiment 2. The design was similar to that of Experiment 1, except for the following changes: (1) only the church was used as the global landmark, and (2) the locations of the fire extinguishers and water fountains were adjusted to ensure that only a subset of them were vertically aligned with the target rooms.

### 3.1 Results and Discussion

A 4 (visual access: no visual access, icon-model, wireframe-model, and window-access)  $\times$  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. 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 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 visualizations 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 conveyed by the two AR models. Without access to 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.

A repeated-measures ANOVA was conducted for each of the three dependent measures of navigation accuracy, navigation 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  $ps > .05$ ). The average drilling accuracy ( $M = 57.4\%$ ,  $SE = 1.9\%$ ) was significantly lower than that found in Experiment 1 ( $M = 89.8\%$ ,  $SE = 1.9\%$ ),  $t(510) = 8.935$ ,  $p < .0001$ , supporting our assertion that the design of the buildings in Experiment 1 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. It appears that

accurate between-floor alignment information, needed in the drilling task, 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 3.

## 4 Experiment 3

The AR visualization models used in Experiment 2 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 2 findings and acknowledging these limitations, we redesigned an X-ray visualization technique in Experiment 3 by allowing navigators to see transparent walls, the global landmark, and the horizon of the outdoor space, as shown in Fig. 5. The X-ray visualization provides 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 3. In addition, users could turn on/off the AR information on-demand.



**Fig. 5.** An X-ray visualization with depth information

A second goal of Experiment 3 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 [20]. Visual

access to multiple global landmarks has been used to help self-localization in outdoor spaces (see [17] 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 3, we evaluated the X-ray visualization with two global landmark conditions (single global landmark access vs. multiple global landmarks access), compared to two control conditions (no visual access to outdoor spaces vs. direct window-access), as were used in Experiment 2. 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. Sixteen new students participated in Experiment 3. The design was the same as Experiment 2, except that only one visualization interface was evaluated but with two global landmark conditions.

#### 4.1 Results and Discussion

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 being more accurate than the window-access condition and no visual access condition (all  $ps < .0005$ ). Interestingly, the results demonstrated that the X-ray visualization (single global landmark access) outperformed the gold standard of window-access in promoting users' development of multi-level cognitive maps. We interpret this superior pointing 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. 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, 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) ( $p = .284$ ). 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 not surprising for two reasons: first, users only required one global landmark (the church) for self-localization in the current studies. Second, 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.

A repeated-measures ANOVA was conducted for each of the three dependent measures of navigation accuracy, navigation 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 navigation accuracy,  $F(3, 189) = .539, p = .656, \eta^2 = .014$ ; or navigation efficiency,  $F(3, 189) = .550, p = .649, \eta^2 = .009$ . These results are consistent with the earlier two experiments. This lack of 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 [14]. The virtual multi-level building in that study had two elevators and the results showed that a sparsely rendered model significantly promoted users' navigation accuracy and efficiency. Thus, we predict that the X-ray visualization used in Experiment 3 could also promote users' wayfinding performance in a complex building with multiple elevators, which will be the topic of a future experiment.

Of note, a significant main effect of visual access was observed for the drilling task,  $F(3, 189) = 5.548, p = .001, \eta^2 = .081$ . Subsequent pairwise comparisons showed that the 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 ( $p = .039$ ). This was not surprising for the no visual access condition but very meaningful for the window-access condition. The finding that the X-ray visualization outperformed the window-access condition in promoting users' drilling accuracy suggests that this interface is a more than adequate substitute for the gold standard of windows. As predicted, it provided clearer inter-floor visualization than was possible from the windows. Taken together, the results of Experiment 3 provide compelling evidence that the X-ray visualization is an effective approach for promoting users' development of a multi-level cognitive map.

## 5 General Discussion

The primary goal of this work was to investigate whether visual access to a global landmark from within a multi-level building, either through direct window access or AR technology, could help multi-level cognitive map development. A multi-level cognitive map represents the globally coherent mental representation of a multi-story built environment. It was evaluated in the current studies using three cross-level spatial tasks including pointing and wayfinding between floors and a cross-floor drilling task.

The most important finding from Experiment 1 is that increasing visual access to a global landmark (both indoor and outdoor) through direct window access significantly promotes users' development of a multi-level cognitive map. 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. 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. Previous literature has discussed how increasing visual access to important level-related building features such as elevators could support users' spatial learning and wayfinding of multi-level buildings [14, 15]. Our current research extends these earlier studies and demonstrates that increasing visual access to a global indoor or outdoor landmark can also facilitate the development of a multi-level cognitive map. This research also provides new insights into our understanding of the underlying mental processes involved in the integration of multi-level spatial knowledge into a multi-level cognitive map; 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 navigators have direct or indirect (via AR visualization) access to a global landmark.

On the basis of the Experiment 1 findings, we proposed and evaluated three AR interfaces in Experiments 2 and 3 (an icon-model, a wireframe-model and an X-ray visualization), compared to two control conditions. The results of Experiment 2 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. The most important finding from Experiment 3 is that the X-ray visualization was not only effective but actually outperformed the "gold standard" of window-access in promoting users' development of multi-level cognitive maps. 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 multi-level cognitive maps. 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.

Taken together, the findings of these experiments provide three Human-computer interaction principles for cognitively motivated visualization techniques for development of indoor navigation systems: first, designers should provide the depth information of 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 on demand rather than having them constantly expose to this information.

**Acknowledgements.** This research was supported by NSF grant CHS-1425337 and NIH grant R01-EY019924-07.

## References

1. Tolman, E.C.: Cognitive maps in rats and men. *Psychol. Rev.* **55**, 189–208 (1948)
2. O'Keefe, J., Nadel, L.: *The hippocampus as a cognitive map*. Oxford University Press, Oxford (1978)

3. Li, H., Giudice, N.A.: Using mobile 3D visualization techniques to facilitate multi-level cognitive map development of complex indoor spaces. In: Graf, C., Giudice, N.A., Schmid, F. (eds.) *Proceedings of the International Workshop on Spatial Knowledge Acquisition with Limited Information Displays, SKALID 2012, Monastery Seon, Germany*, pp. 31–36, August 2012
4. Li, H., Giudice, N.A.: 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. ACM, Orlando (2013)
5. Vidal, M., Berthoz, A.: Navigating in a virtual 3D maze: body and gravity, two possible reference frames for perceiving and memorizing. *Spat. Cogn. Comput.* **5**, 139–161 (2005)
6. Yartsev, M.M., Ulanovsky, N.: Representation of three-dimensional space in the hippocampus of flying bats. *Science* **340**, 367–372 (2013)
7. Jeffery, K.J., Jovalekic, A., Verriotis, M., Hayman, R.: Navigating in a three-dimensional world. *Behav. Brain Sci.* **36**, 523–587 (2013)
8. Finkelstein, A., Derdikman, D., Rubin, A., Foerster, J.N., Las, L., Ulanovsky, N.: Three-dimensional head-direction coding in the bat brain. *Nature* **517**, 159–164 (2014)
9. Hölscher, C., Büchner, S., Strube, G.: Multi-floor buildings and human wayfinding cognition. *Behav. Brain Sci.* **36**, 551–552 (2013)
10. Klatzky, R.L., Giudice, N.A.: The planar mosaic fails to account for spatially directed action. *Behav. Brain Sci.* **36**, 554–555 (2013)
11. Wang, R.F., Street, W.N.: What counts as the evidence for three-dimensional and four-dimensional spatial representations? *Behav. Brain Sci.* **36**, 567–568 (2013)
12. Garling, T., Böök, A., Lindberg, E., Arce, C.: Is elevation encoded in cognitive maps? *J. Environ. Psychol.* **10**, 341–351 (1990)
13. Tlauka, M., Wilson, P.N., Adams, M., Souter, C., Young, A.H.: An investigation into vertical bias effects. *Spat. Cogn. Comput.* **7**, 365–391 (2007)
14. Giudice, N.A., Li, H.: The effects of visual granularity on indoor spatial learning assisted by mobile 3D information displays. In: Stachniss, C., Schill, K., Uttal, D. (eds.) *Spatial Cognition 2012*. LNCS, vol. 7463, pp. 163–172. Springer, Heidelberg (2012)
15. Hölscher, C., Meilinger, T., Vrachliotis, G., Brösamle, M., Knauff, M.: Up the down staircase: wayfinding strategies in multi-level buildings. *J. Environ. Psychol.* **26**, 284–299 (2006)
16. Carlson, L.A., Hölscher, C., Shipley, T., Conroy, D.R.: Getting lost in buildings. *Curr. Dir. Psychol. Sci.* **19**, 284–289 (2010)
17. Steck, S.D., Mallot, H.A.: The role of global and local landmarks in virtual environment navigation. *Presence Teleoperators Virtual Environ.* **9**, 69–83 (2000)
18. Giudice, N.A., Walton, L.A., Worboys, M.: The informatics of indoor and outdoor space: a research agenda. In: *2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*, pp. 47–53 (2010)
19. Dey, A., Sandor, C.: Lessons learned: evaluating visualizations for occluded objects in handheld augmented reality. *Int. J. Hum Comput Stud.* **72**, 704–716 (2014)
20. Loomis, J.M., Klatzky, R.L., Golledge, R.G., Philbeck, J.W.: Human navigation by path integration. *Wayfinding Behav. Cogn. Mapp. other Spat. Process* 125–151 (1999)